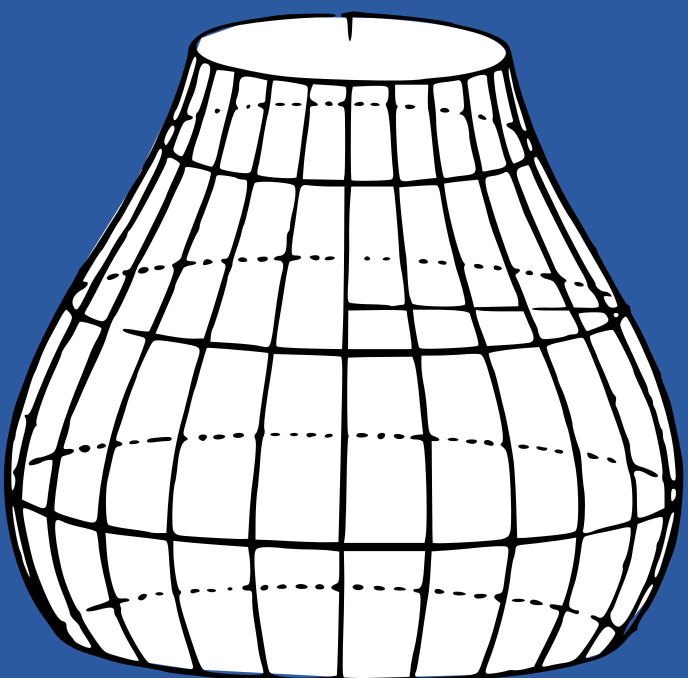
N.P. Tarasov

A Course of Advanced Mathematics for Technical Schools



A COURSE OF ADVANCED MATHEMATICS FOR TECHNICAL SCHOOLS

ı t

A COURSE OF ADVANCED MATHEMATICS FOR TECHNICAL SCHOOLS

N. P. TARASOV

Translated by

D. E. BROWN, M.A.

English Translation Editor

B. G. WALKER, M.Sc.

Department of Mathematics

Brunel College of Technology

Acton, London

PERGAMON PRESS

OXFORD · LONDON · NEW YORK · PARIS

1961

1961

Pergamon Press Ltd.

A translation of the original volume

Н. П. Тарасов

Курс Высшей Математики для техникумов

Library of Congress Card Number 61-10944

Printed in Poland to the order of Państwowe Wydawnictwo Naukowe,
by Drukarnia Uniwersytetu im. Adama Mickiewicza Poznań

CONTENTS

PART II

ELEMENTS OF DIFFERENTIAL CALCULUS

ADDENDUM

PREFACE TO THE NINTH EDITION

This present edition differs only in details from the previous one. Revision has been confined to correcting some inaccuracies and slightly awkward passages in the eighth edition and to simplifying the working in the articles dealing with Fourier series.

I must thank Professor M. F. Bokshtein and Mr. D. M. Toidze for numerous valuable suggestions during the preparation for the press of this ninth edition of *A Course of Higher Mathematics.* I am also grateful to Mr. M. M. Goryachi, the editor of Gostekhizdat, whose careful work has been of considerable assistance in improving some points of the exposition.

N. TARASOV

PREFACE TO THE EIGHTH EDITION

THE present eighth edition of *A Course of Higher Mathematics* differs considerably from earlier editions. The two main reasons for the revision have been the need to bring the text-book into line with the new programme for Technical School mathematics and the desire to simplify and improve certain sections of the course. I have been guided during the revision not only by the requirements of the new programme but also by criticisms from teachers and by the instructions of the pedagogic department of the Ministry of Education board for semispecialist schools and colleges.

The chapters on "Theory of limits" and "Indefinite integrals" have been rewritten for the present edition, whilst the chapters on "The straight line" and "Definite integrals" have undergone considerable revision. The "Foreword" to the book has been rewritten and extended, whilst the general "Conclusion" has been added to the main body of the text. The "Addendum" includes a short section on differential equations and a chapter on "Series" (numerical, power and Fourier series). In line with the requirements of the programme and teachers' recommendations, the sections "Parametric equations of curves" and "The elementary functions and their graphs" have been cut from the book and the treatment shortened for the section "Continuous functions" and for some other less important topics; the chapter "Polar co-ordinates" has been omitted from the addendum.

I have received a great deal of valuable advice whilst working on the book from Prof. M. F. Bokshtein and Mr. A. A. Zykov, to whom I wish to express my indebtedness. I must also thank Mr. V. A. Solodkov, the editor of Gostekhizdat, for his scrupulous attention and corrections to the manuscript.

N. TARASOV

FOREWORD TO THE RUSSIAN EDITION

THE subjects that make up a school course in mathematics, i.e. arithmetic, algebra, geometry and trigonometry, share the common name of "elementary mathematics". Whilst a knowledge of elementary mathematics is required for an understanding of such mathematical disciplines as are treated in the present book (analytic geometry, differential and integral calculus), these latter form at the same time the basis of "higher mathematics".

Elementary mathematics is concerned, at bottom, with a study of *constant* quantities and figures. The solution of algebraic equations, for instance, consisting as it does in seeking the roots, i.e. constant quantities, may be regarded as a typical problem of elementary algebra. Geometry is concerned with studying the properties of fixed geometric figures. The main attention in trigonometry is focused on the trigonometric transformations and evaluation of the elements of triangles.

Mathematics is a tool which men can use to explore and modify their natural environment. But it can never become an effective tool if it is to be confined to the study of *constant* quantities, since nature is made up essentially of ceaseless change. Thus mathematics only became the basis of the natural and applied sciences after systematic methods had been found for investigating *variable* quantities—so that the relationships could be studied between the variables that participate in natural phenomena and productive processes.

The problem of acquiring a mastery over variables was particularly acute for the mathematicians of the 16th and 17th centuries, when capitalism was beginning its rapid expansion in the wake of feudalism. The rise of capitalism was accompanied by an enormous increase in production, which led in turn to rapid growth of the natural and applied sciences. Practical requirements faced scientists, and in particular mathematicians, with a series of new problems demanding urgent solution.

It was soon seen that these new problems needed an entirely new *mathematical apparatus*, which would allow real natural phenomena to be studied whilst *in the process of change*. Such an apparatus was in fact created in the 17th century by the combined efforts of mathematicians of various nationalities.

The new apparatus was based on the introduction of variables into mathematics and consisted of *analytic geometry* on the one hand and *mathematical analysis* on the other, the latter subject being primarily made up of the differential and integral calculus.

"The Cartesian variable was the turning point in mathematics", wrote F. Engels. "Thanks to this, *motion* and *dialectic* came into mathematics, and *the differential and integral calculus* became necessary at the precise moment of their perfection, if not invention, by Newton and Leibniz". (F. Engels, *Natural dialectic*, Gospolizdat, 1952, p. 206.)*

Engels observes in the above passage that the differential and integral calculus were *perfected* though not *discovered* (in the second half of the 17th century) by Newton (1642–1727) and Leibniz (1646–1716). The methods of the differential and integral calculus were utilized by scholars long before the works of these two great men; they can be seen in an embryonic form as early as Archimedes (287–212 B.C.), the famous mathematician, physicist and engineer of classical times. The merit of Newton and Leibniz lies, however, in their full appreciation of the profound *internal connexion* between the differential and integral calculus. The idea of logically exploiting this connexion led to a synthesis of the two theories into a single "analysis of infinitesimals". †

---

* The name "Cartesian variable" derives from that of the great French mathematician P. Descartes (1596–1650), who developed the new method of co-ordinates for solving geometrical problems and thus created a new branch of geometry, to be known later as "analytic geometry". The reader will discover in Chapter II of this book that variables appear in geometry as the so-called current co-ordinates of points.

† The student must appreciate that the nature of this connexion will only become clear after studying the differential and integral calculus.

The following words of Engels underline the vital significance of the powerful movement that placed the study of variable quantities at the heart of the mathematical sciences and their applications:

"It seems doubtful if any of the achievements of theoretical science can be reckoned as great a triumph of the human intellect as the invention of the infinitesimal calculus in the second half of the 17th century". (F. Engels, *Natural Dialectic*, Gospolitizdat, 1952, p. 214.)

The logical formulation of the infinitesimal calculus in the works of Newton and Leibniz was followed by a brilliant epoch in the evolution of mathematics. The achievements due to the infinitesimal calculus were inevitably accompanied by an overall advance which embraced, in particular, theoretical mechanics and mathematical physics. The new weapon became what it has since remained, an indispensable part of the equipment of the applied mathematician.

The new discipline led to the development of others. The differential and integral calculus form the foundation at the present time of the imposing edifice known as "mathematical analysis". Russian scientists have played an important role in building up this edifice.

A brief foreword cannot hope to convey a full picture of the vast Russian contribution to mathematics. We shall therefore confine ourselves to mentioning a few of the leading Russian figures in the history of the subject.

The name of Nikolai Ivanovich Lobachevskii (1792–1856) is familiar to serious students throughout the world. Lobachevskii gained immortality by his discovery of a new type of geometry. which differs essentially from that known from the time of Euclid. Indeed Lobachevskii's ideas were so much in advance of his time that they were not understood by many of the leading mathematicians of the day and had to wait for general recognition until after his death.

Lobachevskii also left behind some elegant and valuable studies in the field of mathematical analysis. For instance, he was the first to formulate a general definition of the fundamental analytic concept of function.

A mention of the history of analysis is bound to recall the great Mikhail Vasil'evich Ostrogradskii (1801–1862), a member of the

Russian Academy. A number of Ostrogradskii's results appear in all the modern text-books and his name is familiar to mathematicians of all nations. Apart from his mathematical work, Ostrogradskii also left behind valuable contributions to mechanics and other allied fields.

Pafnutii L'vovich Chebyshev (1821–1894) was another mathematician whose work has gained international recognition whilst leaving its mark on practically every branch of the subject. He was responsible for first class work in applied as well as pure mathematics, as witness his classical researches into the theory of mechanisms.

Chebyshev's work as a teacher was of supreme importance. His pupils included a number of famous men, who helped Chebyshev to found the St. Petersburg school. This latter had an enormous influence on the spread of pure and applied mathematics throughout Russia.

One of Chebyshev's most distinguished pupils was Aleksandr Mikhaillovich Lyapunov (1857–1918), whose studies embraced several branches of mathematics and mechanics. His work is notabe for its accuracy and rigour and includes, in particular, development of the theory of dynamic stability, which is of great importance in applied science.

Sof'ya Vasil'evna Kovalevskaya (1850–1891) was another great Russian mathematician of the last century. Probably her greatest achievement lay in the field of partial differential equations, a subject of vital importance in applied mathematics. Kovalevskaya's theorem on systems of differential equations is now included in all text-books on partial differential equations. She was awarded the prize of the Paris Academy of Sciences in 1888 for her work on the motion of a rigid body about a fixed point. This work gained her an international reputation.

The great Russian mathematicians of the last century worked either in complete isolation, as in the case of Lobachevskii, or within a narrow circle of men of science. Their works remained the property of a handful of specialists. Official government circles were hostile, or at best indifferent, to the spread of science in Russia. Reading

the biographies of these great savants, one can only wonder at the strength of mind with which they defended the inteiests of science against the sluggish inertia of government bodies. The reason is not far to seek, why the upsurge now to be seen in the sciences, and especially in mathematics, began with the October revolution.

The new socialist order aims at maximum satisfaction of the material and cultural requirements of society as a whole. Party and Government alike, therefore, pay the fullest possible attention to the advancement of science. Our country has been covered with a vast network of establishments for higher education and scientific research institutions. Education is available to every citizen of the Soviet Union. Science has become public property. The period of the solitary savant of Tsarist days has been replaced by one in which scientists combine in powerful associations to conquer together the difficulties inherent in creative endeavour.

In addition to the old scientific and mathematical centres (Moscow, Leningrad, Kazan, Kharkov), new centres have arisen since the October revolution, at Tbilisi, Saratov, Tashkent, Kiev, Odessa, Gorki, Tomsk, Sverdlovsk and other cities.

The fact of whole mathematical associations working on the same problem has led to the formation of Soviet mathematical *schools*, headed by famous Soviet scientists.

An outstanding iole in the development of mathematics has been played by the theory of functions school, closely associated with the name of Nikolai Nikolaevich Luzin (1883–1950). The work of Luzin and his pupils M. Ya. Suslin (1894–1919), P. S. Aleksandrov (born 1896), A. N. Kolmogorov (born 1903), A. Ya. Khinchin (born 1894), D. E. Menshov (born 1892) has brought the theory of functions school to the forefront of world science.

The theory of probability school, continuing in the direction laid down by Chebyshev and his pupils Lyapunov and A. A. Markov (1856–1922), can claim quite exceptional achievements. The works of A. N. Kolmogorov, A. Ya. Khinchin, S. N. Bernshtein (born 1880), N. V. Smirnov (born 1900) and other Soviet mathematicians have been decisive in the development of probability theory and its applications.

The theory of numbers school is headed by I. M. Vinogradov (born 1891), a member of the Academy and Hero of Socialist Labour. His researches have produced a whole new trend in number theory.

It is impossible to give a brief mention here to the great Soviet achievements in other departments of mathematics. We shall merely recall the important role in the development of the theory of partial differential equations played by I. G. Petrovskii (born 1901) and S. L. Sobolev (born 1908).

A good many Soviet mathematicians have by no means confined their attention to a single branch of the subject. Kolmogorov, for instance, is one of the most versatile and penetrating of modern mathematicians.

Kolmogorov and other Soviet workers have applied mathematical theory to the solution of purely industrial problems with outstanding success, thus carrying on the tradition of Russian science whereby theory is never divorced from practice but is rather made to guide the latter. Here is what Chebyshev had to say about the connexion between theory and practice: "The confrontation of theory with practice yields extremely valuable results, and it is not only practice which benefits; the pure sciences themselves are advanced by an interplay which reveals new subjects for scientific study or new facets in long familiar subjects". (*Izbr. Matemat. Trudy* (Selected Mathematical Works), Gostekhizdat, 1946, p. 100.)

Impressive examples of the union of theory with practice are provided by the work of such leading engineers as Zhukovskii (1847–1921), Chaplygin (1869–1942) and Krylov (1863–1945).

Zhukovskii is one of the founders of aeronautical science; in fact, he is often called "the father of Russian aviation". His discoveries in this field were not merely significant in his own time. In particular, he demonstrated theoretically the possibility of the figures of advanced pilotage—theoretical work that was impressively vindicated by Captain P. N. Nesterov, the first in the world to "loop the loop". Zhukovskii was also concerned with numerous other technical and mechanical problems, which were equally distinguished by a high mathematical level.

Chaplygin, one of Zhukovskii's pupils, was possibly the most remarkable mechanical engineer of recent times. He used new mathematical methods with outstanding success in the theory of aeronautics. He was likewise interested in purely mathematical problems and gave, for instance, a new method for the approximate solution of certain classes of differential equation (later perfected by Luzin).

Krylov's work in the fields of ship-building and navigation has become classical; and, like Zhukovskii and Chaplygin, he left behind him several valuable mathematical studies. Krylov was above all an expert on approximation methods and designed a complex mathematical computer. He also paid considerable attention to the history of mathematics and translated Newton's *Mathematical Foundations of Natural Philosophy*, which he supplied with valuable commentaries.

PART I

THE ELEMENTS OF PLANE ANALYTIC GEOMETRY

CHAPTER 1

THE RECTANGULAR CO-ORDINATES OF POINTS ON A PLANE. ELEMENTARY APPLICATIONS OF THE CO-ORDINATE METHOD

§ **1. Rectangular system of co-ordinates.** We fix the unit of length (scale) and initial point O for measurements on a straight line. The point O splits the line into two rays; we shall call the direction of one ray positive, and of the other, negative.

A straight line on which a measurement origin O and scale are fixed and a positive direction chosen is called a number axis. This designation is bound up with the method well known to readers of representing real numbers with the aid of a straight line. A number axis is alternatively called a *co-ordinate axis* or *axis of co-ordinates.*

Two mutually perpendicular co-ordinate axes with a common origin O (at their point of intersection) *form a rectangular plane Cartesian * co-ordinate system* (fig. 1). We shall always assume

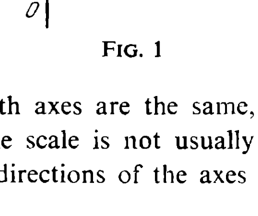


FIG. 1

hat the scales on both axes are the same, though this is not in general necessary. The scale is not usually specially indicated on figures. The positive directions of the axes are defined by arrows.

* After the great French mathematician Descartes cf. footnote on p. 14.

One of the co-ordinate axes is called the axis of *abscissae* (axis Ox or x-axis) and the other the axis of *ordinates* (axis Oy or y-axis).

The origin divides each axis into two parts, positive (in the positive direction from the origin) and negative (in the negative direction from the origin).

The axes will always be assumed to be mutually situated so that the shortest rotation of the positive part of the axis of abscissae about the origin to make it coincide with the positive part of the axis of ordinates is anti-clockwise.

§ **2. Rectangular co-ordinates of a point in a plane.** We take a point M of the plane xOy lying on neither of the axes Ox, Oy and draw perpendiculars MN and MP from this point to the axis of abscissae and axis of ordinates respectively. The bases of the perpendiculars, points N and P, will lie either on the positive or on the negative parts of the corresponding axes (fig. 2).

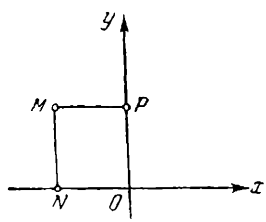


FIG. 2

The *abscissa* of the point M is defined as the length of ON, taken with a plus sign if N lies on the positive part of the axis of abscissae and with a minus sign if N is on the negative part. The abscissa of the point M is usually denoted by the letter x.

The *ordinate* of M is the length of OP, taken with a plus sign if P lies on the positive part of the axis of ordinates, and with a minus sign if P is on the negative part. The ordinate of M is usually denoted by the letter y.

In accordance with these definitions, the abscissa of the point M shown in fig. 2 is negative, and the ordinate positive.

If M lies on the axis Ox, its ordinate y is considered zero; if M lies on the axis Oy, its abscissa x is considered zero. Hence

if M coincides with the origin O, its abscissa and ordinate are both zero.

The abscissa x and ordinate y of a point M on the plane are called the *rectangular Cartesian co-ordinates* of M. We shall refer to them in future simply as the *co-ordinates* of M.

Obviously, a unique pair of co-ordinates x, y corresponds to each point M of the plane.

Conversely, a unique point M of the plane corresponds to each pair of numbers x, y, these numbers being the co-ordinates of the point. In fact, let both numbers x and y differ from zero. We measure a distance ON equal to $|x|$* from the origin O in the positive direction of Ox if $x>0$, and in the negative direction if $x<0$. Similarly we measure a distance OP equal to $|y|$ in the positive direction of Oy if $y>0$, and in the negative direction if $y<0$. From the ends N and P we erect perpendiculars to the axes Ox and Oy respectively; these perpendiculars intersect in a unique point M of the plane, the numbers x and y being clearly the co-ordinates of M.

If $x\neq 0$ and $y=0$, the point M corresponding to these numbers will lie on the x-axis, since the distance OP is now zero, whereas if $x=0$ and $y\neq 0$, we obtain a unique point M lying on the y-axis. Finally, if $x=0$ and $y=0$, the corresponding point M coincides with the origin O.

Thus a unique pair of co-ordinates x, y corresponds to each point M of the plane, and conversely, a unique point M of the plane corresponds to each pair of co-ordinates x, y. The point M is said to be completely defined by its co-ordinates.

When a point M has co-ordinates x, y we write $M(x, y)$.

It follows from all that has been said that the introduction of the concept of the co-ordinates of a point enables problems concerning the positions of points on a plane to be solved with the aid of numbers.

The co-ordinate axes divide the total plane into four parts called *quadrants*, numbered as shown in fig. 3.

---

* The symbol $|x|$ denotes the absolute value of the number x cf. p. 123.

The signs of the co-ordinates of points lying in the respective quadrants are shown in the table below:

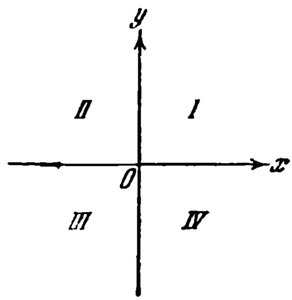


FIG. 3

Quadrant	*Sign of abscissa*	*Sign of ordinate*
I	+	+
II	−	+
III	−	−
IV	+	−

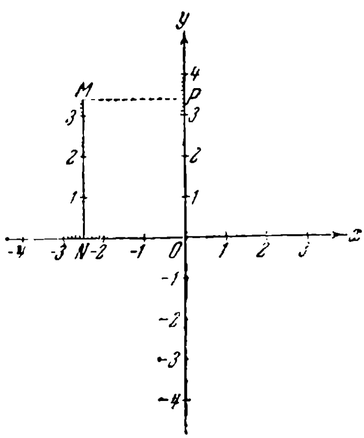


FIG. 4

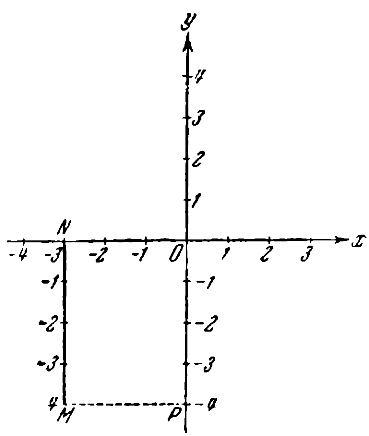


FIG. 5

Example 1. Given a point M on the plane (fig.4). To find its abscissa and ordinate (the position of the co-ordinate axes is assumed known).

Solution. From M we drop perpendiculars MN and MP on to the axes Ox and Oy respectively and measure the distances ON and OP *; we find that $x=-ON=-2{\cdot}5$, $y=NM=3{\cdot}4$. We arrive at the same result on dropping the perpendicular MN from M on to the axis Ox and measuring the distances ON and MN.

Example 2. To construct the point $M(-3, -4)$, given the co-ordinate axes.

Solution. We measure a distance ON equal to three units of length along the axis Ox to the left of the origin O (noting that the abscissa of M is negative). We erect a perpendicular to the axis Ox from the end N and measure downwards along this a distance equal to four units of length (the ordinate of the required point is negative). The end of this segment is in fact the required point M (fig. 5). We get the same point M on measuring distances $ON=|-3|$ and $OP=|-4|$ along the negative x- and y-axes respectively and drawing straight lines through N and P parallel to the axes to intersect in the point M.

§ 3. **The distance between two points.** 1. Given the points $M_1(x_1, y_1)$ and $M_2(x_2, y_2)$, it is required to find the distance between them, *i.e.* the length of the straight line $M_1 M_2$.

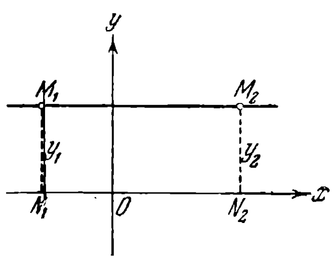


FIG. 6

We use the same notation M_1M_2 to denote both the actual line and the length of the line joining M_1 and M_2.

Suppose that the line joining the points M_1, M_2 is parallel to the x-axis (*i.e.* that $y_1=y_2$). Figure 6 illustrates the case when

* The scale is assumed to be indicated on both axes.

one end (M_1) has a negative, and the other end (M_2) has a positive abscissa. We have from fig. 6:

$$M_1 M_2 = N_1 N_2 = N_1 O + O N_2 .$$

Since here $x_1 < 0$ (which means that $-x_1 > 0$), $N_1 O = -x_1$*. whilst $ON_2 = x_2$, so that

$$M_1 M_2 = -x_1 + x_2 = x_2 - x_1 .$$

If M_1 and M_2 change places, the abscissa x_2 becomes negative and x_1 positive, the difference $x_2 - x_1$ becomes negative, and hence the distance $M_1 M_2$ (a positive quantity) is given by the number $-(x_2 - x_1)$, which is clearly equal to the absolute value of the difference $x_2 - x_1$. Thus we have, independently of whether the difference $x_2 - x_1$ is a positive or negative number:

$$M_1 M_2 = |x_2 - x_1| . \tag{1}$$

It may easily be seen (we leave this to the reader) that expression (1) remains true whenever the distance $M_1 M_2$ lies along a line parallel to the x-axis (or along the x-axis itself).

If $x_1 = x_2$, i.e. if $M_1 M_2$ lies along line parallel to the y-axis (or on the y-axis itself), we find similarly that

$$M_1 M_2 = |y_2 - y_1| . \tag{2}$$

2. If $x_2 \neq x_1$ and $y_2 \neq y_1$, the straight line passing through M_1, M_2 is not parallel to a co-ordinate axis (fig. 7).

Through M_1, M_2 we draw straight lines parallel to the axes Ox and Oy respectively, and denote their point of intersection by R; the co-ordinates of R will be (x_2, y_1), as may readily be seen. We find from expressions (1), (2):

$$M_1 R = |x_2 - x_1| , \quad RM_2 = |y_2 - y_1| .$$

We have from the right-angled triangle $M_1 RM_2$:

$$M_1 M_2 = \sqrt{(M_1 R)^2 + (RM_2)^2} ;$$

* Since the abscissa x_1 is negative, its absolute value is $-x_1$. By definition of abscissa (§ 2) this absolute value is equal to the distance $N_1 O$.

on substituting in this for $(M_1 R)^2$ and $(RM_2)^2$ in terms of the co-ordinates we finally get

$$M_1 M_2 = \sqrt{(x_2 - x_1)^2 + (y_2 - y_1)^2}\,. \tag{3}$$

Instead of writing the sum of the squares of the absolute values of differences $x_2 - x_1$ and $y_2 - y_1$ under the radical in equation

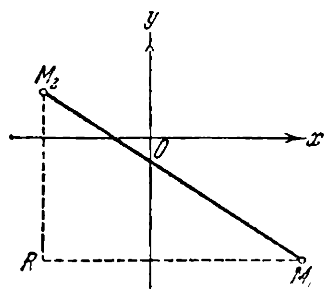


FIG. 7

(3), we write the sum of the squares of the differences themselves since the square of the absolute value of a number is equal to the square of the number itself.

The square root in (3) is taken with the + sign, since the distance between two points is a positive number. This number will be denoted by the letter *d*.

By using (3), we can easily find the distance d between a point $M(x, y)$ and the origin: we set $x_1 = y_1 = 0$, and $x_2 = x$, $y_2 = y$, and obtain:

$$d = \sqrt{x^2 + y^2}\,.$$

Example. To find the point equidistant from three given points $M_1(1, 2)$, $M_2(-1, -2)$, $M_3(2, -5)$.

Solution. To find the point means to find its co-ordinates. Let M be the required point, with co-ordinates x, y. To find the two unknowns x and y we require two equations, which are obtained by using the conditions given in the problem. The two equations are

$$M_1 M = M_2 M \text{ and } M_2 M = M_3 M\,.$$

We have from expression (3):

$$M_1 M = \sqrt{(x-1)^2 + (y-2)^2},$$

$$M_2 M = \sqrt{(x+1)^2 + (y+2)^2},$$

$$M_3 M = \sqrt{(x-2)^2 + (y+5)^2}.$$

Thus we get two equations in the unknowns x and y:

$$\sqrt{(x-1)^2 + (y-2)^2} = \sqrt{(x+1)^2 + (y+2)^2},$$

$$\sqrt{(x+1)^2 + (y+2)^2} = \sqrt{(x-2)^2 + (y+5)^2}.$$

We find on solving these equations: $x = \frac{8}{3}$, $y = -\frac{4}{3}$. Thus the required point is $M\left(\frac{8}{3}, -\frac{4}{3}\right)$.

§ 4. **Division of a distance in a given ratio.** This implies the following problem:

We are given the two points $M_1(x_1, y_1)$ and $M_2(x_2, y_2)$ (fig. 8).

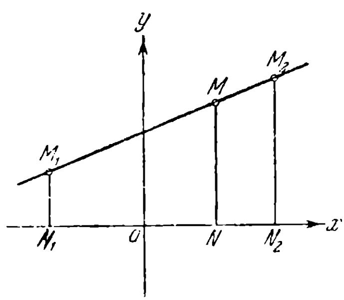


FIG. 8

A third point M, whose co-ordinates are unknown, divides the distance $M_1 M_2$ in such a way that the ratio $\frac{M_1 M}{M M_2}$ is equal to a given number λ. We require to find the point M, i.e. its co-ordinates x, y.

We know from elementary geometry that the intercepts M_1M, MM_2, N_1N, NN_2 formed by parallels are proportional. Hence on taking into account the given condition we can write:

$$\frac{M_1M}{MM_2}=\frac{N_1N}{NN_2}=\lambda .$$

But we have from expression (1):

$$N_1N=|x-x_1|, \quad NN_2=|x_2-x| .$$

Since the required point M lies between M_1 and M_2 on the line joining them, the differences $x-x_1$ and x_2-x are simultaneously both positive or both negative for any positions of M_1 and M_2. Hence the ratio

$$\frac{x-x_1}{x_2-x}$$

is always positive, and we can take this instead of the ratio

$$\frac{|x-x_1|}{|x_2-x|} .$$

We can thus write:

$$\frac{M_1M}{MM_2}=\frac{N_1N}{NN_2}=\frac{x-x_1}{x_2-x}=\lambda ,$$

whence we find for the abscissa x of M:

$$x=\frac{x_1+\lambda x_2}{1+\lambda} . \tag{4}$$

We obtain the ordinate y of M in a similar manner:

$$y=\frac{y_1+\lambda y_2}{1+\lambda} . \tag{5}$$

In particular, if M bisects M_1M_2, we have $\lambda=\dfrac{M_1M}{MM_2}=1$, and aequtions (4) and (5) become:

$$x=\frac{x_1+x_2}{2}, \tag{4*}$$

$$y=\frac{y_1+y_2}{2}. \tag{5*}$$

Example 1. To find the point $M(x, y)$ dividing the line joining $M_1(2, 3)$ and $M_2(3, -3)$ in the ratio $\frac{2}{5}$.

Solution. In this case $\lambda=\frac{2}{5}$, $x_1=2$, $x_2=3$, $y_1=3$, $y_2=-3$, and expressions (4) and (5) give;

$$x=\frac{2+\frac{2}{5}\cdot 3}{1+\frac{2}{5}}=\frac{16}{7}, \quad y=\frac{3+\frac{2}{5}\cdot(-3)}{1+\frac{2}{5}}=\frac{9}{7}.$$

Thus the required point is $M\left(\frac{16}{7}, \frac{9}{7}\right)$.

Example 2. To show that the straight line joining the mid-points of two sides of a triangle is parallel to the third side.

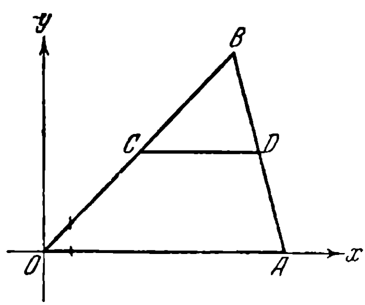


FIG. 9

Solution. We arrange the co-ordinate axes so that the origin coincides with one vertex of the triangle whilst the x-axis coincides with a side passing through this vertex (fig. 9). The vertices are now $O(0, 0)$, $A(x_1, 0)$, $B(x_2, y_2)$. The straight line CD joins the

mid-points of sides OB, AB. We want to show that CD is parallel to OA, i.e. parallel to the x-axis, for which it is sufficient to show that the ordinates of points C, D are equal. The ordinates y_C, y_D of C and D are given by expression (5*):

$$y_C=\frac{0+y_2}{2}=\frac{y_2}{2}, \quad y_D=\frac{0+y_2}{2}=\frac{y_2}{2}.$$

Hence $y_C=y_D$, i.e. CD is parallel to the side OA.

Example 3. Two point-masses at $M_1(x_1, y_1)$, $M_2(x_2, y_2)$ with masses m_1, m_2 respectively are acted on by gravity. Find the centre of gravity of the system.

Solution. We know from mechanics that the centre of gravity lies on $M_1 M_2$ at a point $M(x, y)$ that divides $M_1 M_2$ in the inverse ratio to the gravitational forces acting on M_1, M_2, i.e. in the ratio $\lambda=\frac{m_2 g}{m_1 g}=\frac{m_2}{m_1}$, where g is the acceleration due to gravity. On taking this into account, we find from expressions (4) and (5):

$$x=\frac{x_1+\frac{m_2}{m_1}x_2}{1+\frac{m_2}{m_1}}, \quad y=\frac{y_1+\frac{m_2}{m_1}y_2}{1+\frac{m_2}{m_1}},$$

or

$$x=\frac{m_1 x_1+m_2 x_2}{m_1+m_2}, \quad y=\frac{m_1 y_1+m_2 y_2}{m_1+m_2}.$$

EXERCISES

1. Construct the points with the co-ordinates:

$$x=3,\ y=5;\ x=-2,\ y=0;\ x=2,\ y=-4;$$

$$x=0,\ y=3;\ x=3,\ y=-3;\ x=\sqrt{2},\ y=1.$$

2. Given the point $x=4$, $y=-3$, find the co-ordinates of the point symmetrical to it with respect to (*i*) the axis of abscissae, (*ii*) the axis of ordinates.

Ans. $(4, 3)$; $(-4, -3)$.

3. Find the point symmetrical to the point $(3, -5)$ with respect to the bisector of the first quadrant.

Ans. $(-5, 3)$.

4. Given a square with a side equal to 3 units of length, find the vertices of the square when any two of its non-parallel sides are taken as co-ordinate axes.

Ans. $(0, 0)$; $(3, 0)$; $(3, 3)$; $(0, 3)$;
or $(0, 0)$; $(0, 3)$; $(-3, 3)$; $(-3, 0)$;
or $(0, 0)$; $(-3, 0)$; $(-3, -3)$; $(0, -3)$;
or $(0, 0)$; $(0, -3)$; $(3, -3)$; $(3, 0)$.

5. Given a square with a side equal to 4 units of length, find the co-ordinates of the vertices of the square if its diagonals are taken as axes.

Ans. $(2\sqrt{2}, 0)$; $(0, 2\sqrt{2})$; $(-2\sqrt{2}, 0)$; $(0, -2\sqrt{2})$.

6. Given a rhombus with a side equal to 5 units of length and one diagonal equal to 6 units, find the co-ordinates of the vertices of the rhombus if its diagonals are taken as co-ordinate axes.

Ans. $(3, 0)$; $(0, 4)$; $(-3, 0)$; $(0, -4)$;
or $(4, 0)$; $(0, 3)$; $(-4, 0)$; $(0, -3)$.

7. Find the co-ordinates of the vertices of a regular hexagon of side a when the origin is located at the centre of the hexagon and the axis of abscissae passes through two opposite vertices.

Ans. $(a, 0)$; $\left(\frac{a}{2}, \frac{a\sqrt{3}}{2}\right)$; $\left(-\frac{a}{2}, \frac{a\sqrt{3}}{2}\right)$;
$(-a, 0)$; $\left(-\frac{a}{2}, -\frac{a\sqrt{3}}{2}\right)$; $\left(\frac{a}{2}, -\frac{a\sqrt{3}}{2}\right)$.

8. Find the perimeter of a triangle with vertices $(3, 4)$, $(-2, 4)$, $(2, 2)$.

Ans. $5+3\sqrt{5}$.

9. Find the perimeter of a quadrilateral with vertices $(2, 1)$, $(-2, 8)$, $(-6, 5)$, $(-2, -2)$.

Ans. $10+2\sqrt{65}$.

10. Find the perimeter of a quadrilateral with vertices $(-a, 0)$, $(0, b)$, $(a, 0)$, $(0, -b)$.

Ans. $4\sqrt{a^2+b^2}$.

11. Show that the triangle with vertices $(-3, -2)$, $(1, 4)$, $(-5, 0)$ is isosceles.

12. Show that the triangle with vertices $(-1, 1)$, $(1, 3)$, $(-\sqrt{3}, 2+\sqrt{3})$ is equilateral.

13. Show that the triangle with vertices $(1, 2)$, $(3, 4)$, $(-1, 4)$ is right-angled.

14. Show that the points $(8, 0)$, $(0, -6)$, $(7, -7)$, $(1, 1)$ lie on a circle whose centre is $(4, -3)$. What is the radius of the circle?

Ans. 5.

15. Find the point equidistant from the points $(0, 0)$, $(1, 0)$, $(0, 2)$.

Ans. $\left(\frac{1}{2}, 1\right)$.

16. Find the point equidistant from the points $(-4, 3)$, $(4, 2)$, $(1, -1)$.

Ans. $\left(\frac{1}{18}, \frac{17}{18}\right)$.

17. Find the point equidistant from the points $(1, 3)$, $(0, 6)$, $(-4, 1)$.

Ans. $\left(-\frac{73}{34}, \frac{123}{34}\right)$.

18. Find the centre of the circle passing through the points $(0, 0)$, $(4, 2)$, $(6, 4)$.

Ans. $(-3, 11)$.

19. Find the point on the Ox axis equidistant from the points $(0, 5)$ and $(4, 2)$.

Ans. $\left(-\frac{5}{8}, 0\right)$.

20. Find the points simultaneously 5 units of length distant from the point $(1, 3)$ and 4 units of length from the y-axis.

Ans. $(4, 7)$, $(4, -1)$, $(-4, 3)$.

21. A point moves along a straight line from the point $A(-3, -2)$ to $B(4, 5)$. Find the length of the path traversed and the angle φ that the path makes with the axis of abscissae.

Ans. $AB=\sqrt{98}$, $\phi=45°$.

22. Find the positions of points that start from $A(3, 2)$ and travel 12 units of length along a straight line forming an angle of 60° with the axis of abscissae.

Ans. $[9, 2(1+3\sqrt{3})]$, $[-3, 2(1-3\sqrt{3})]$.

23. Find the point dividing the distance between points $P_1(-2, 3)$ and $P_2(4, 6)$ in the ratio 2:3.

Ans. $\left(\frac{2}{5}, \frac{21}{5}\right)$.

24. One end of a straight line is at $(2, 5)$ and its mid-point is at $(-1, 2)$. Find the co-ordinates of the other end.

Ans. $(-4, -1)$.

25. The point $(1, 1)$ bisects the straight line joining points $(x, 5)$ and $(-2, y)$. Find these points.

Ans. $(4, 5)$, $(-2, -3)$.

26. Find the point dividing the straight line joining points $(0, 2)$ and $(8, 0)$ in a ratio equal to the ratio of the distances of these points from the origin.

Ans. $\left(\frac{8}{5}, \frac{8}{5}\right)$.

27. Divide the distance between the points $(-3, 4)$ and $(9, 12)$ in the same ratio as that of the distances of these points from the origin.

Ans. $(0, 6)$.

28. Two vertices of a triangle are at $A(3, 7)$ and $B(-2, 5)$. Find the co-ordinates of the third vertex, given that the mid-points of the sides passing through it lie on the co-ordinate axes.

Ans. $(-3, -5)$ or $(2, -7)$.

29. Given the points $A(1, -1)$ and $B(4, 5)$, to what point must AB be produced (in the direction from A to B) in order that its distance from A is three times the length AB?

Ans. $(10, 17)$.

30. Find the lengths of the medians of a triangle whose vertices are $A(3, 4)$, $B(-1, 1)$, $C(0, -3)$.

Ans. $\frac{5}{2}\sqrt{5}$, $\frac{1}{2}\sqrt{26}$, $\frac{1}{2}\sqrt{149}$.

31. If a triangle has vertices $P_1(1, 2)$, $P_2(3, -4)$, $P_3(5, 5)$, find the point of intersection of its medians.

Ans. $(3, 1)$.

32. Find the points trisecting the distance between points $M_1(-3, -7)$ and $M_2(10, 2)$.

Ans. $\left(\frac{4}{3}, -4\right)$, $\left(\frac{17}{3}, -1\right)$.

33. If a triangle has vertices $(5, 0)$, $(3, -8)$, $(1, -4)$, find the points trisecting its medians.

Ans. $(3, -4)$, $(4, -2)$, $(3, -6)$, $(2, -4)$.

34. Show that in any right-angled triangle the length of the median joining the mid-point of the hypotenuse to the opposite vertex is equal to half the length of the hypotenuse.

35. In the trapezium $OABC$ the parallel sides OA, CB are perpendicular to the side OC. If D is the mid-point of AB, show that $OD=CD$.

36. Weights of 60 and 40 grammes are located at the points $A(4, 6)$ and $B(-2, 7)$ respectively. Find the co-ordinates of the centre of gravity of the system.

Ans. (1.6, 6.4).

37. Weights of 30, 50 and 70 grammes are located at the points $A(-1, 0)$, $B(-2, 4)$, $C(4, -5)$ respectively. Find the centre of gravity of the system.

Ans. $(1, -1)$.

Hint. First find the centre of gravity M of the system consisting of any two of the weights, say those at A and B, then find the centre of gravity of the weights at M and C.

38. Show that the system made up of n masses $m_1, m_2, \ldots, m_n$ concentrated at the points $A_1(x_1, y_1)$, $A_2(x_2, y_2)$, ..., $A_n(x_n, y_n)$ respectively has a centre of gravity whose co-ordinates are given by

$$x=\frac{x_1 m_1+x_2 m_2+\ldots+x_n m_n}{m_1+m_2+\ldots+m_n}$$

$$y=\frac{y_1 m_1+y_2 m_2+\ldots+y_n m_n}{m_1+m_2+\ldots+m_n}.$$

39. Given that the centre of gravity of a homogeneous triangular plate lies at the point of intersection of the medians, find the co-ordinates of the centre of gravity in terms of the co-ordinates (x_1, y_1), (x_2, y_2), (x_3, y_3) of the vertices.

Ans. $x=\frac{x_1+x_2+x_3}{3}$, $y=\frac{y_1+y_2+y_3}{3}$.

CHAPTER 2

THE STRAIGHT LINE

§ 5. **The straight line as the geometrical locus of points equidistant from two given points. The equation of a straight line. General equation of a straight line.** We considered in the previous chapter the question of defining the position of a point on a plane with the aid of numbers. Some simple problems were also solved, the solution of which rested entirely upon the application of this idea.

The co-ordinate method is not merely applied, however, to problems relating to the positions of individual points. It happens that an extension of this basic idea—the definition of the position of a point on a plane with the aid of co-ordinates—enables more complex geometrical shapes—curves—to be studied by algebraic methods. Our discussion of the subject will start with a consideration of the simplest type of curve, i.e. the straight line.

1. We know from elementary geometry that *every* point equidistant from the two points A_1, A_2 lies on a straight line PQ which is the perpendicular bisector of A_1A_2. The line PQ is said to be the *geometrical locus of points equidistant from A_1 and A_2.*

We consider PQ, the geometrical locus of points equidistant from $A_1(2, 4)$ and $A_2(6, -2)$ (fig. 10).

The condition that a point $M(x, y)$ lies on PQ is expressed by the equation

$$A_1M = A_2M. \tag{1}$$

From expression (3) § 3, defining the distance between two points in terms of their co-ordinates, we have

$$A_1M = \sqrt{(x-2)^2+(y-4)^2},$$

$$A_2M = \sqrt{(x-6)^2+(y+2)^2}.$$

We find on substituting these expressions in equation (1):

$$\sqrt{(x-2)^2+(y-4)^2}=\sqrt{(x-6)^2+(y+2)^2}\,.$$

We have obtained the relationship between the co-ordinates (x, y) of the point M expressing the condition that M lies on the straight line PQ.

Relationship (2) can be written in a simpler form. By squaring both sides, removing the brackets, taking the right-hand side over to the left, collecting like terms and cancelling we obtain

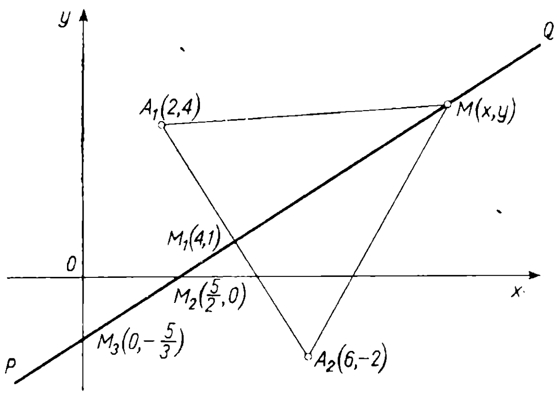


FIG. 10

$$x^2-4x+4+y^2-8y+16=x^2-12x+36+y^2+4y+4\,,$$

$$8x-12y-20=0\,,$$

or finally, on dividing through by 4, we arrive at

$$2x-3y-5=0\,. \tag{3}$$

2. When a point $M(x, y)$ moves on a plane its co-ordinates x and y vary. Since the co-ordinates of a moving point are variables they are referred to as *current co-ordinates*. Suppose we regard the x and y in relationship (3) as current co-ordinates. Then the relationship becomes an equation in two variables.

If the current co-ordinates x, y in equation (3) are replaced by numbers, instead of the expression $2x-3y-5$ on the left-hand side

we get a certain number; if the replacement leads to the equation becoming the identity $0=0$, the numbers substituted for x and y are said to *satisfy* equation (3), whereas if the left-hand side gives a non-zero number, we say that the numbers substituted for x and y *do not satisfy* equation (3) (in this case (3) no longer becomes the identity $0=0$).

Equation (3) expresses the condition that the point M lies on the straight line PQ. It means that the co-ordinates of every point M lying on the straight line must satisfy the equation of the line, whilst the co-ordinates of any point not on the straight line cannot satisfy its equation (3).

For instance, the point $M_1(x_1, y_1)$ lying at the centre of A_1A_2 is on our straight line; from expressions (4*) and (5*) of § 4, the co-ordinates of M_1 are

$$x_1=\frac{2+6}{2}=4\,, \qquad y_1=\frac{4-2}{2}=1\,.$$

On substituting the numbers 4 and 1 respectively for the current co-ordinates x, y in equation (3), the left-hand side vanishes:

$$2.4-3.1-5=0\,,$$

and (3) becomes the identity $0=0$.

We now take a point not on PQ, say $A_1(2\,,\,4)$; on replacing the current co-ordinates x, y in equation (3) by the numbers 2 and 4 respectively, we obtain on the left-hand side the number

$$2.2-3.4-5=-13$$

instead of zero. Hence the equation does not reduce to $0=0$ as a result of the substitution.

Equation (3) is thus *satisfied by the co-ordinates of any point lying on the straight line PQ, and is not satisfied by the co-ordinates of any point not on this straight line*. We therefore refer to equation (3) as *the equation of the straight line PQ*.

The above leads us to the following conclusions:

(a) In order to decide whether a given point lies on PQ or not, it is sufficient to see whether or not its co-ordinates satisfy equation

(3); if they do, the point lies on PQ; whereas if they do not, then the point is not on PQ.

For example, the point $N_1(7, 3)$ lies on PQ, since

$$2.7-3.3-5=0;$$

whereas $N_2(-1, 3)$ is not on PQ, since

$$2.(-1)-3.3-5\neq 0.$$

(b) Given equation (3) of the straight line PQ, we can find as many points on PQ as we wish. For suppose one of the current co-ordinates, say x, is assigned some arbitrary numerical value. On replacing x by this number (3) becomes an equation with respect to y. On solving this latter equation for y, we get a pair of numbers (co-ordinates) satisfying the equation of the straight line, or in other words, the co-ordinates of a point lying on the straight line.

On setting $x=-2$, for instance, we get

$$2(-2)-3y-5=0,$$

whence

$$y=-3;$$

so that the point $A(-2, -3)$ lies on the straight line.

Let us now find the co-ordinates (x, y) of the point of intersection of PQ with the x-axis. We know one of the current co-ordinates, i.e. ordinate $y=0$. We substitute zero for y in equation (3) and get

$$2x-5=0,$$

whence

$$x=\frac{5}{2}.$$

Hence the required point is $\left(\frac{5}{2}, 0\right)$ (M_2 in fig. 10).

Similarly, to find the co-ordinates of the point of intersection M_3 of PQ with the y-axis, we write zero for the abscissa x in equation (3) and find the value of the co-ordinate y from the equation

$$-3y-5=0;$$

we get $y=-\frac{5}{3}$, so that M_3 has co-ordinates $\left(0, -\frac{5}{3}\right)$.

3. We found in 2 of this section the equation of a straight line as the geometrical locus of points equidistant from points $A_1\,(2\,,4)$ and $A_2(6\,,-2)$. We now consider the general form of the problem.

Let the straight line PQ be given as the locus of points equidistant from $A_1(a_1, b_1)$ and $A_2(a_2, b_2)$*. The condition that $M\,(x, y)$ lies on PQ is expressed by the equation

$$A_1 M = A_2 M\,. \tag{4}$$

As in the above particular case, we find the equation of the straight line by expressing the distances A_1M and A_2M in equation (4) in terms of the co-ordinates of A_1, A_2, and M.

We have from the formula for the distance between two points

$$A_1 M = \sqrt{(x-a_1)^2+(y-b_1)^2}\,,$$

$$A_2 M = \sqrt{(x-a_2)^2+(y-b_2)^2}$$

and consequently

$$\sqrt{(x-a_1)^2+(y-b_1)^2} = \sqrt{(x-a_2)^2+(y-b_2)^2}\,.$$

This is in fact the required equation of the straight line. Like equation (4), it expresses the condition that the point $M\,(x\,,y)$ lies on the straight line.

We can simplify equation (5) by squaring both sides and removing the brackets; it then becomes

$$\begin{aligned} x^2-2a_1\,x+a_1^2+y^2-2b_1\,y+b_1^2 \\ = x^2-2a_2\,x+a_2^2+y^2-2b_2\,y+b_2^2\,. \end{aligned}$$

After taking the right-hand side over to the left and collecting and cancelling like terms, we get

$$2\,(a_2-a_1)\,x+2\,(b_2-b_1)\,y+a_1^2+b_1^2-a_2^2-b_2^2=0\,.$$

For brevity we write

$$2\,(a_2-a_1)=A\,,\quad 2\,(b_2-b_1)=B\,,\quad a_1^2+b_1^2-a_2^2-b_2^2=C\,. \tag{6}$$

---

* We naturally exclude the case when simultaneously $a_1=a_2$ and $b_1=b_2$, since the points A_1 and A_2 are then coincident.

The equation of the straight line forming the locus of points equidistant from the points $A_1(a_1, b_1)$ and $A_2(a_2, b_2)$ thus finally becomes

$$Ax+By+C=0 .$$

Clearly, everything said in 2 regarding equation (3) holds good for equation (7).*

3. Equation (7) is an equation of the first degree in the current co-ordinates (x, y) of a point of the straight line. Since every straight line can be regarded as the locus of points equidistant from some two given points, in addition to finding equation (7) for a straight line we have also arrived at the following theorem:

THEOREM. *Any straight line can be expressed by a first degree equation in the current co-ordinates x, y* (i.e. the co-ordinates of an arbitrary point of the line).

We refer to (7) as *the general equation of a straight line.*

§ 6. The equations of straight lines parallel to the co-ordinate axes. The equations of the co-ordinate axes. 1. We take the straight line PQ parallel to the y-axis and intersecting the x-axis in the point $K(l, 0)$ (fig. 11). With these conditions the position of the straight line in the plane (relative to the co-ordinate system xOy) is fully defined.*

We can regard the given line PQ as the locus of points equidistant from the origin O and the point $N(2l, 0)$.

The condition that a point $M(x, y)$ lies on PQ is expressed by

$$OM=NM .$$

---

* In deducing equation (7) we removed the radicals in equation (5) by squaring both sides. We know that such an operation may lead to an equation not equivalent to the original, i.e. to an equation which is satisfied not only by the co-ordinates satisfying equation (5) but also by other ("redundant") co-ordinates. It may be readily shown that in the present case equation (7) is satisfied only by the co-ordinates that satisfy equation (5). The proof of this assertion will be omitted, however, both now and in future similar cases.

* If l is a positive number, the straight line lies to the right of the y-axis and if negative, to the left.

We have from the formula for the distance between two points:

$$OM=\sqrt{x^2+y^2}\,, \qquad NM=\sqrt{(x-2l)^2+y^2}\,,$$

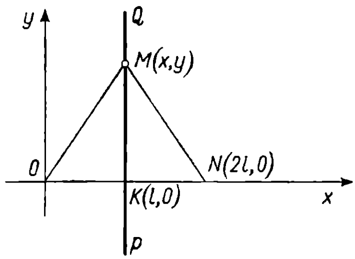


FIG. 11

and since $OM=NM$, we have the following equation:

$$\sqrt{x^2+y^2}=\sqrt{(x-2l)^2+y^2}\,.$$

Squaring both sides of this equation gives us

$$x^2+y^2=x^2-4lx+4l^2+y^2\,,$$

or on simplifying:

$$-x4l+4l^2=0\,,$$

whence, on dividing through by 41 and changing all the signs, we obtain

$$x-l=0\,. \tag{7}$$

Equation (8) is a particular case of the general equation

$$Ax+By+C=0\,. \tag{8}$$

Equation (8) is obtained from (7) with $A=1$, $B=0$, and $C=-l$. This means that we can write equation (8) in full as

$$x+0\cdot y-l=0\,. \tag{9}$$

Since the product of any number and zero is zero, it follows from equation (9) that for any value of the ordinate y the abscissa x of a point $M(x\,,y)$ on PQ satisfies the equation

$$x-l=0 \text{ or } x=l\,,$$

and this is precisely the characteristic property of all points of a line parallel to the Oy axis and cutting the Ox axis in the point $K(l, 0)$.

The absence in equation (8) of a term containing the current co-ordinate y is explained by the vanishing of the coefficient of y in this equation.

2. Suppose now we are given a straight line parallel to the x-axis and intersecting the y-axis in the point $K(0, l)$ (fig. 12). This

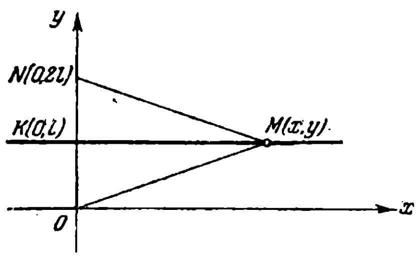


FIG. 12

straight line is clearly the locus of points equidistant from the points $O(0, 0)$ and $N(0, 2l)$. The condition that the point $M(x, y)$ lies on the straight line is expressed by

$$OM = NM .$$

We obtain the equation of the straight line by expressing the lengths of OM and NM in terms of the co-ordinates of their ends

$$\sqrt{x^2+y^2} = \sqrt{x^2+(y-2l)^2} .$$

On squaring both sides of this equation and simplifying, we obtain in turn

$$x^2+y^2 = x^2+y^2-4ly+4l^2 ,$$

$$4ly-4l^2=0$$

or finally:

$$y-l=0 . \tag{10}$$

Equation (10) is obtained from the general equation (7) with $A=0$, $B=l$ and $C=-l$. Thus the absence in equation (10) of a

term containing the current co-ordinate x is explained by the vanishing of the coefficient of x in this equation. Equation (10) expresses the characteristic property of all points on a line parallel to the x-axis, namely that for any abscissa x the ordinate y remains equal to l.

3. With $l=0$ the straight line parallel to Oy becomes the y-axis itself. Consequently, the equation of the y-axis is

$$x=0\,. \tag{11}$$

Similarly, the equation of the x-axis is

$$y=0\,. \tag{12}$$

Equation (11) indicates that for all values of the ordinate y the abscissa x of points on the straight line remains zero. Thts is precisely the property of every point of the y-axis.

Equation (12) shows that for all values of the abscissa x the ordinate y of points on the straight line remains zero. This is precisely the property of every point of the x-axis.

§ 7. **The equation of an inclined straight line.** 1. We define *the angle of inclination of a straight line with the axis Ox* as the angle

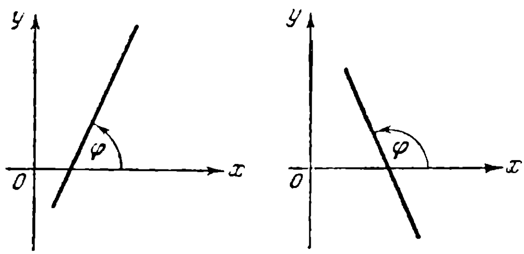


FIG. 13

through which the positive direction of the x-axis must be turned, about its point of intersection with the straight line, in order to make it coincide with the straight line (fig. 13).

If the straight line is parallel to the x-axis, its angle of inclination to Ox is considered *zero.* On writing φ for the angle of inclination of a straight line with the x-axis, we always have $0 \leq \varphi < 180°$.

2. A straight line (not parallel to Oy) is fully defined by specifying (*i*) its point of intersection $(0, b)$ with Oy, and (*ii*) its angle of inclination φ with Ox.

Let a straight line be specified by these two conditions, the angle of inclination φ being different from zero.

As we saw in § 5, to find the equation of a given straight line we have to establish a geometrical equality expressing the condition that a point $M(x, y)$ lies on the straight line and then write this equality in co-ordinate form (in terms of the current co-ordinates

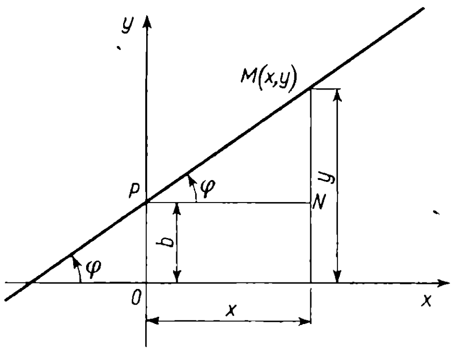


FIG. 14

x, y of M). The relationship thus obtained is the equation of the given straight line.

As is readily established from fig. 14, the condition that $M(x, y)$ lies on the given straight line is expressed by the equation

$$NM = PN \tan\varphi . \tag{13}$$

We now write this equation in co-ordinate form; since $NM = y - b$, $PN = x$, (13) becomes

$$y - b = x \tan\varphi . \tag{13*}$$

The tangent of the angle of inclination φ of a straight line with the axis Ox is called the slope of the straight line and is usually denoted by the letter k. On writing k for $\tan\varphi$ in equation (13*)

and taking the term $-b$ to the right-hand side, we get what may be called *the slope equation of the straight line:*

$$y = kx + b\,. \tag{14}$$

The ordinate b of the point of intersection of the straight line with Oy is called the *initial ordinate.*

Equation (14) was deduced on the assumption that the angle of inclination φ of the straight line with Ox differs from zero. But equation (14) can also represent a straight line for which $\varphi = 0°$, i.e. a line parallel to Ox. For if $\varphi = 0°$, $k = \tan 0° = 0$ and equation (14) takes the form

$$y = b\,.$$

This is the equation of a straight line parallel to Ox for it has the same form as equation (10), obtained earlier from other considerations (see § 6).

A straight line perpendicular to Ox does not intersect Oy and has an angle of inclination of 90°. For such a straight line there exists neither an initial ordinate b nor a slope k (tan 90° does not exist). Consequently, a straight line perpendicular to Ox cannot be expressed by a slope equation.

To represent a straight line by a slope equation we need to know its slope and initial ordinate. Hence the problem of finding the slope equation of a straight line consists in finding the coefficients k and b of equation (14) in accordance with the data given.

Example 1. A straight line is inclined to Ox at an angle of 30° and has an initial ordinate equal to -3. Find the equation of the straight line.

Solution. We have by hypothesis: $k = \tan 30° = \frac{1}{\sqrt{3}}$; $b = -3$.

On substituting these values in equation (14) we get the required equation of the straight line as $y = \left(\frac{1}{\sqrt{3}}\right)x - 3$.

Example 2. Find the equation of the straight line passing through the point $(2\,, -5)$ and inclined at 45° to the axis Ox.

Solution. Finding the equation $y=kx+b$ amounts to finding the two unknowns k and b. To do this, the conditions of the problem must yield two equations containing the unknowns.

Since we are told that the straight line is inclined at 45° to Ox, we obtain the first equation

$$k=\tan 45°=1\ ,$$

which happens in this case to be solved with respect to the unknown k.

We know in addition that the required straight line passes through the point $(2\ ,\ -5)$ and this means that the co-ordinates $(2\ ,\ -5)$ must satisfy the required equation. This gives

$$-5=2k+b\ .$$

We have thus obtained the two equations:

$$\left.\begin{aligned} k&=1\ , \\ -5&=2k+b\ . \end{aligned}\right\}$$

Hence $k=1$ and $b=-7$ and the required equation is therefore

$$y=x-7\ .$$

Note. We often speak simply of "finding a straight line," rather than of "finding the equation of a straight line."

§ 8. THEOREM: **Every first-degree equation in the current co-ordinates represents a straight line.** We showed in 3 in § 5, that any straight line can be expressed by a first-degree equation in the current co-ordinates. We now prove the converse.

THEOREM. *Every first-degree equation in the current co-ordinates, i.e. every equation of the form*

$$Ax+By+C=0\ , \tag{15}$$

represent a straight line. (We assume that at least one of the coefficients A, B is non-zero, for if $A=0$, $B=0$, we should no longer have an equation in x and y).

We first assume that $B \neq 0$; on solving equation (15) for y, it becomes

$$y = -\frac{A}{B}x - \frac{C}{B}. \tag{15*}$$

We now consider the straight line whose slope $k = -\frac{A}{B}$ and whose initial ordinate $b = -\frac{C}{B}$. As we know, this straight line is represented by the equation

$$y = kx + b,$$

i.e. by equation (15*). Equation (15*) thus represents a straight line; and since this equation is equivalent to equation (15), (15) also represents a straight line.

We now take the case $B = 0$. Equation (15) now becomes $Ax + C = 0$, or

$$x + \frac{C}{A} = 0. \tag{15**}$$

We consider the straight line parallel to Oy that intersects Ox in the point with abscissa $l = -\frac{C}{A}$. We know from § 7 that this straight line is given by the equation $x - l = 0$, i.e. by (15**). But since this equation is equivalent to equation (15), this latter equation also represents a straight line.

Thus, whatever the coefficients A, B, C (provided that at least one of A, B is non-zero), equation (15) represents a straight line, and the theorem is proved.

§ 9. **Incomplete equations of straight lines.** We have just shown that every first-degree equation

$$Ax + By + C = 0 \tag{15}$$

represents a straight line. We consider the particular cases when one or two of the coefficients A, B, C vanish.

I. $C=0,\ A\neq 0,\ B\neq 0$.

Equation (15) becomes

$$Ax+By=0\,. \tag{16}$$

The co-ordinates of the point $(0\,,0)$, i.e. the origin, satisfy equation (16). Hence, *if the free term is absent in the equation of a straight line, the straight line passes through the origin.*

II. $A=0,\ B\neq 0,\ C\neq 0$.

Equation (15) becomes $By+C=0$ or

$$y=-\frac{C}{B}, \tag{17}$$

and this is the equation of a straight line parallel to Ox (cf. § 6).

III. $B=0,\ A\neq 0,\ C\neq 0$.

In this case (15) becomes $Ax+C=0$ or $x=-\dfrac{C}{A}$, and represents a straight line parallel to Oy (cf. § 6).

It follows from cases II and III that, *if the equation of a straight line contains no term in x, the straight line is parallel to Ox; whilst if there is no term in y, the straight line is parallel to Oy.*

IV. $A=0,\ C=0,\ B\neq 0$.

Equation (15) becomes $By=0$ or $y=0$; as we saw in § 6, this is the equation of the x-axis.

V. $B=0,\ C=0,\ A\neq 0$.

Equation (15) becomes $Ax=0$ or $x=0$; as we saw in § 6, this is the equation of the y-axis.

Example 1. The straight line $3x+5y=0$ passes through the origin, since the free term in the equation is zero.

The straight line $2x-5=0$ is parallel to Oy, since the term in y is absent.

The equation $3x=0$ can be reduced to $x=0$ and therefore represents the y-axis.

Example 2. To construct the straight line given by

$$3x+4y-11=0.$$

Solution. We first find any two points lying on the straight line by the method of 2*b*. in § 5. We put $x=1$ "say" in the given equation. We find for y the value $y=2$. On putting $x=-3$, we get $y=5$. We have thus found two points $(1, 2)$ and $(-3, 5)$ on the straight line. We plot these points and draw the straight line

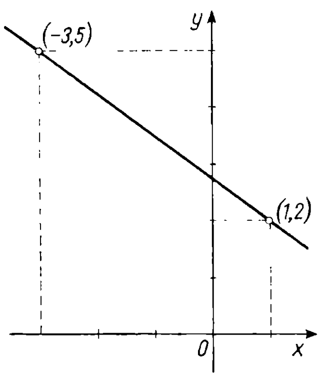


FIG. 15

through them (fig. 15); this is the straight line defined by the given equation.

§ 10. Equation of a straight line passing through a given point (the equation of a pencil of straight lines). We want to find the equation of a straight line passing through a given point $M(x_1, y_1)$. We shall seek the equation in the form (14):

$$y=kx+b,$$

where k and b are coefficients to be defined. k and b are often called parameters.

The single condition contained in the problem—that the straight line pass through the given point—is not sufficient to define the two unknown coefficients. Evidently, all we can do in this case is express one parameter in terms of the other, and thus arrive

at an equation containing an arbitrary coefficient, i.e. a coefficient that can take any numerical values. This means that we must obtain an equation that represents an infinite set of straight lines instead of a single line. This situation does not contradict the given hypothesis since an infinite set of straight lines can in fact pass through a given point.

We turn to the solution.

Since the point $M(x_1, y_1)$ lies on the straight line, the co-ordinates (x_1, y_1) must satisfy equation (14), i.e.

$$y_1 = kx_1 + b.$$

We use this equation to write b in terms of k:

$$b = y_1 - kx_1.$$

On substituting this value for b in equation (14), we get

$$y = kx + y_1 - kx_1 \text{ or}$$

$$y - y_1 = k(x - x_1). \tag{18}$$

The slope k in this equation remains indeterminate and can therefore take different values. For the different values of k we obtain the equations of different straight lines. But all these straight lines pass through the point $M(x_1, y_1)$, since the co-ordinates x_1, y_1 clearly satisfy equation (18) for any value of k. Hence we say that equation (18) is *the equation of the pencil of straight lines passing through the point* $M(x_1, y_1)$.

Example. To find the equation of the pencil of straight lines passing through the point $(-3, 5)$.

Solution. On substituting $x_1 = -3$, $y_1 = 5$ in equation (18), we get

$$y - 5 = k(x + 3),$$

i.e. the equation of the required pencil, where k is an arbitrary factor.

§ 11. Equation of the straight line passing through two given points. Let the two given points be $M_1(x_1, y_1)$ and $M_2(x_2, y_2)$.

where $x_1 \neq x_2$, $y_1 \neq y_2$. We want to find the equation of the straight line passing through these two points.

We first write the equation of the pencil of straight lines passing through one of the given points, say $M_1(x_1, y_1)$. This equation has the form (18):

$$y - y_1 = k(x - x_1).$$

The required straight line is the one belonging to pencil (18) passing through the point $M_2(x_2, y_2)$. This means that the k in equation (18) must have a definite value for the required equation, viz. the value for which (18) is satisfied by the co-ordinates x_2, y_2 of the point M_2. We obtain the required value of k by substituting x_2, y_2 for the current co-ordinates x, y in (18):

$$k = \frac{y_2 - y_1}{x_2 - x_1}.$$

The required equation will thus be

$$y - y_1 = \frac{y_2 - y_1}{x_2 - x_1}(x - x_1),$$

or

$$\frac{y - y_1}{y_2 - y_1} = \frac{x - x_1}{x_2 - x_1}. \tag{19}$$

Example 1. Find the equation of the straight line passing through the points $(1, 2)$ and $(-3, 5)$.

Solution. We get by substituting $x_1 = 1$, $y_1 = 2$, $x_2 = -3$, $y_2 = 5$ in equation (19):

$$\frac{y - 2}{5 - 2} = \frac{x - 1}{-3 - 1},$$

or

$$3x + 4y - 11 = 0.$$

We can check our solution simply by verifying that the co-ordinates of the given points satisfy the equation obtained. We obtain on substituting the co-ordinates of one point then of the other for the current co-ordinates in the equation:

$$3.1+4.2-11=0\,,$$

$$3(-3)+4.5-11=0$$

and see that the solution is in fact correct.

Note. We deduced equation (19) on the assumption that $x_1 \neq x_2$ and $y_1 \neq y_2$. If $x_1 = x_2$ or $y_1 = y_2$, one of the denominators in equation (19) vanishes; and since division by zero is impossible, the straight line cannot be represented by equation (19). If $x_1 = x_2$, the straight line passing through points M_1, M_2 is parallel to Oy whilst if $y_1 = y_2$, the straight line is parallel to Ox. But these straight lines are represented by equations (8) and (10) respectively (§ 6) and there is therefore no need to use equation (19) in these cases.

§ **12. Intercept equation of a straight line.** When we are given that a straight line passes through the points $A(a\,,0)$ and $B(0\,,b)$, i.e. when we are given the points of intersection of the straight line with the axes, the equation obtained for the straight

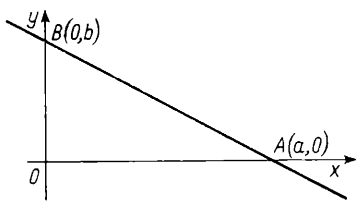


FIG. 16

line is known as its *intercept equation* (we assume that $a \neq 0$ and $b \neq 0$). (See fig. 16).

We are concerned here with a particular case of the specification of a straight line by two points. The required equation is therefore obtained by using equation (19), in which we set $x_1 = a$, $y_1 = 0$, $x_2 = 0$, $y_1 = b$. This gives us

$$\frac{y-0}{b-0} = \frac{x-a}{0-a}\,, \tag{20}$$

which readily reduces to the form

$$\frac{x}{a}+\frac{y}{b}=1\,. \tag{21}$$

This is the final form for the *intercept equation of a straight line.*

Example. Find the equation of the straight line intersecting the axis Ox in the point $(5\,,0)$ and the axis Oy in the point $(0\,,\,-3)$.

Solution. On setting $a=5$, $b=-3$ in equation (21) we get

$$\frac{x}{5}+\frac{y}{-3}=1\,,$$

or

$$3x-5y-15=0\,.$$

§ 13. Angle between two straight lines. 1. Let the intersecting straight lines AB and CD (fig. 17) be given by the equations

$$y=k_1x+b_1\,,$$

$$y=k_2x+b_1\,.$$

We shall assume that these lines are not perpendicular.* They form two angles at their point of intersection P: the acute angle $\angle BPD=\alpha$ and the obtuse angle $\angle DPA=180°-\alpha$. For certain problems it becomes necessary to distinguish between these two angles.

We call one of the angles formed by the straight lines the angle between CD and AB, and the other the angle between AB and CD.

We follow the definition of the angle of inclination of a straight line to the Ox axis and call *the angle between CD and AB the angle through which AB must be turned in a positive direction* (i.e. anticlockwise) *about the point P in order to make it coincide with CD.* In fig. 17 this angle is $\angle BPD=\alpha$. The angle between AB and CD is similarly defined. (This angle is in our case $\angle DPA=180°-\alpha$.)

---

* It will be shown in § 14 that the straight lines intersect if $k_1\neq k_2$, and are not perpendicular if $k_1k_2\neq -1$. We assume that the slopes of given lines satisfy both these conditions.

2. We now turn to the problem of finding the angle α between CD and AB, given their equations.

Let AB form an angle φ_1 with the axis Ox, and CD an angle φ_2. From fig. 17 we have $\alpha=\varphi_2-\varphi_1$; hence

$$\tan\alpha=\tan(\varphi_2-\varphi_1)=\frac{\tan\varphi_2-\tan\varphi_1}{1+\tan\varphi_1\tan\varphi_2}.$$

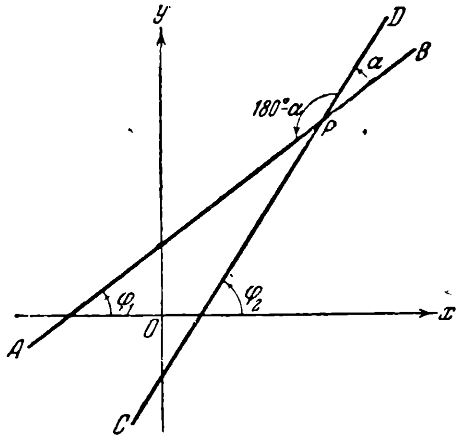


FIG. 17

But we know (§ 7) that $\tan\varphi_1=k_1$, $\tan\varphi_2=k_2$; hence

$$\tan\alpha=\frac{k_2-k_1}{1+k_1k_2}. \tag{22}$$

We deduce the expression for the tangent of the angle formed by AB with CD. Writing β for this angle, we have

$$\tan\beta=\tan(180^\circ-\alpha)=-\tan\alpha=-\frac{k_2-k_1}{1+k_1k_2},$$

or

$$\tan\beta=\frac{k_1-k_2}{1+k_1k_2}. \tag{22*}$$

Expressions (22) and (22*) differ only in sign. It may be noted that the slope subtracted in the numerators of these expressions

belongs to the straight line with which the given straight line forms the angle defined by (22) or (22*).

If a problem gives no indication as to which angle between the straight lines is required, the choice of the straight line from which the angle is considered is of no importance, i.e. it makes no difference whether we use (22) or (22*) for solving the problem.

We explain the above with the aid of examples.

Example 1. Find the angle between the straight lines

$$2x=3y+5=0 \quad \text{and} \quad x+2y+2=0\,.$$

Solution. We solve the equations for y and thus reduce them to the form containing the slope:

$$y=\frac{2}{3}x+\frac{5}{3}\,; \qquad y=-\frac{1}{2}x-1\,.$$

Since the problem does not say whether the required angle is that formed by the first line with the second or conversely it does not matter which slope we take as being subtracted. Thus we can write:

$$\tan\alpha=\frac{\frac{2}{3}-(-\frac{1}{2})}{1+\frac{2}{3}\cdot(-\frac{1}{2})}=\frac{7}{4}\,,$$

whence $\alpha=\text{arc tan}\,\dfrac{7}{4}$.

Example 2. Find the equation of the straight line passing through the point $(-2\,,0)$ and forming an angle arc tan $\dfrac{2}{3}$ with the straight line $3x+4y+6=0$.

Solution. We use the equation for a straight line passing through a given point:

$$y-y_1=k(x-x_1)\,,$$

In our example, $x_1=-2$, $y_1=0$. On substituting these values in the equation we get

$$y=k(x+2)\,.$$

For finding the slope k of the required straight line, we note that the angle mentioned is that which the required straight line forms with the given straight line. The angle $\arctan \frac{2}{3}$ must therefore be reckoned from the given straight line; so that in the numerator of the expression for the tangent of this angle we have to subtract the slope of the given straight line from the slope k of the required line, and not vice versa.

We find the slope of the given straight line from its equation as $-\frac{3}{4}$. Hence

$$\frac{2}{3}=\frac{k-\left(-\frac{3}{4}\right)}{1+k\left(-\frac{3}{4}\right)},$$

so that $k=-\frac{1}{18}$. The required equation is therefore

$$y=-\frac{1}{18}(x+2),$$

or

$$x+18y+2=0.$$

On the other hand, if we wanted to find the equation of the straight line passing through the point $(-2, 0)$ with which the straight line $3x+4y+6=0$ forms the angle $\arctan \frac{2}{3}$, the equation for k would be $\dfrac{2}{3}=\dfrac{-\frac{3}{4}-k}{1-\frac{3}{4}k}$, whence $k=-\frac{17}{6}$ and the required equation becomes

$$y=-\frac{17}{6}(x+2),$$

or

$$17x+6y+34=0.$$

§ 14. Conditions for two straight lines to be parallel or perpendicular. 1. *For the straight lines to be parallel.* If the straight lines are parallel they form the same angle with the Ox axis. Hence their slopes k_1 and k_2 are equal ($k_1 = k_2$).

Conversely, if $k_1 = k_2$, the straight lines have the same angle of inclination to Ox, and they are therefore parallel.

Hence, *the condition for two straight lines to be parallel is that their slopes be equal:*

$$k_1 = k_2 . \tag{23}$$

Note. In establishing the condition for two straight lines to be parallel we have not used expressions (22) and (22*) of the previous section, since these were deduced on the assumption that the straight lines intersect. But these expressions may easily be seen to hold good when AB and CD are parallel.

For the angle α between two parallel straight lines is 0°, so that $\tan \alpha = 0$. On the other hand, it follows from the condition

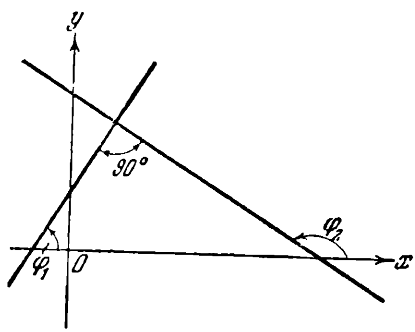


FIG. 18

for the straight lines to be parallel, i.e. from the equation $k_2 = k_1$, that $k_2 - k_1 = 0$, and we get from (22)

$$\frac{k_2 - k_1}{1 + k_2 k_1} = 0 .$$

2. *For the straight lines to be perpendicular.* The angle α between two intersecting straight lines is given by (22) in terms of $\tan \alpha$. If $\alpha = 90°$ (22) is inapplicable since $\tan 90°$ does not exist. This

is why we assumed when deducing (22) that the straight lines did not intersect at right angles.

If the straight lines are perpendicular, we have (fig. 18)

$$\varphi_2 = \varphi_1 + 90^\circ ,$$

whence

$$\tan \varphi_2 = \tan (\varphi_1 + 90^\circ) = -\cot \varphi_1 ,$$

or

$$\tan \varphi_2 = -\frac{1}{\tan \varphi_1} .$$

On replacing $\tan \varphi_1$ and $\tan \varphi_2$ by k_1 and k_2, we find

$$k_2 = -\frac{1}{k_1} ,$$

or

$$1 + k_1 k_2 = 0 .$$

Conversely, let

$$k_2 = -\frac{1}{k_1} .$$

This means that

$$\tan \varphi_2 = -\frac{1}{\tan \varphi_1} ,$$

or

$$\tan \varphi_2 = -\cot \varphi_1 ,$$

or

$$\tan \varphi_2 = \tan (\varphi_1 + 90^\circ) ,$$

whence we obtain:

$$\varphi_2 = \varphi_1 + 90^\circ .$$

The angle between the straight lines is therefore 90°, i.e. the lines are perpendicular.

Hence, *the condition for two straight lines* to be perpendicular s that their slopes have reciprocal absolute values and are opposite in sign:*

$$k_2 = -\frac{1}{k_1} ; \tag{24}$$

* Neither of which is perpendicular to Ox.

Example 1. Find the equation of the straight line parallel to the straight line $3x-5y+6=0$ and passing through the point $(-2, 3)$.

Solution. The equation of the pencil of straight lines passing through the point $(-2, 3)$ has the form (cf. expression (18))

$$y-3=k(x+2).$$

Since the required straight line is parallel to the straight line $3x-5y+6=0$ it must have the same slope, i.e. $\frac{3}{5}$. On substituting $k=\frac{3}{5}$ in the pencil equation we obtain the required equation

$$y-3=\frac{3}{5}(x+2),$$

or

$$3x-5y+21=0.$$

Example 2. Find the equation of the straight line passing through the point $(5, 3)$ and perpendicular to the straight line $7x+9y+1=0$.

Solution. The slope of the given straight line is $-\frac{7}{9}$. Hence by condition (24) the slope of the required straight line is $\frac{9}{7}$. We use expression (18) to write the equation of the pencil of straight lines passing through the point $(5, 3)$ and set $k=\frac{9}{7}$, which gives us the required equation as

$$y-3=\frac{9}{7}(x-5),$$

or

$$9x-7y-24=0.$$

§ **15. Intersection of two straight lines.** Let two straight lines be given by the equations

$$A_1 x + B_1 y + C_1 = 0\,,$$

$$A_2 x + B_2 y + C_2 = 0\,.$$

We want to find the co-ordinates of their point of intersection.

Since the required point lies simultaneously on both the straight lines its co-ordinates must satisfy both the given equations. Thus, *to find the point of intersection of two straight lines, we must solve their equations simultaneously for x and y.*

Example 1. Find the point of intersection of the straight lines

$$5x - y - 7 = 0\,,$$

$$3x + 2y - 12 = 0\,.$$

Solution. On multiplying the first equation by two and adding to the second, we get $13x - 26 = 0$, whence $x = 2$. On substituting this value for x in the first equation, we find $y = 3$. Hence the required point is $(2\,,3)$.

Example 2. Find the point of intersection of the straight lines

$$3x + 4y - 2 = 0\,,$$

$$6x + 8y + 7 = 0\,.$$

Solution. The equations given are incompatible and have no solution.

This fact is readily explained geometrically: the slopes of the straight lines given are the same, i.e. we are concerned with parallel straight lines for which no point of intersection exists.

Example 3. Find the point of intersection of the straight lines

$$2x - y + 1 = 0\,,$$

$$6x - 3y + 3 = 0\,.$$

Solution. We see by inspection that the second equation leads to the first on dividing through by three. Thus the given system reduces to a single equation and has an infinite set of solutions:

we can obtain as many solutions as we wish by assigning an arbitrary value to one unknown and working out the corresponding value of the other.

This is explained geometrically by the fact that the two lines coincide, so that every point of one is at the same time a point of the other.

Such a system of equations is said to be *indeterminate.*

EXERCISES

1. Does the straight line

$$y=3x+13$$

pass through the points: (*a*) $(5, -1)$; (*b*) $(-4, 1)$; (*c*) $(3, 2)$; (*d*) $(5, 7)$; (*e*) $(-5, -2)$?

2. Does the straight line

$$3x-4y+11=0$$

pass through the points: (*a*) $(3, 5)$, (*b*) $(\frac{1}{2}, -1)$, (*c*) $(2, -4)$ (*d*) $(-1, 2)$?

3. Draw the straight lines given by the following equations: (*a*) $y=x$; (*b*) $y=2x$; (*c*) $y=\frac{1}{2}x$; (*d*) $y=2x+3$; (*e*) $y=-x$; (*f*) $y=-2x$; (*g*) $y=-\frac{1}{2}x$; (*h*) $y=-2x-3$; (*i*) $2x+3y-4=0$; (*j*) $2x-3y-4=0$; (*k*) $3x+2=0$; (*l*) $3y-5=0$; (*m*) $2x+3y=0$.

4. Find the equation of the straight line forming an angle of 30° with Ox and meeting Oy in the point $(0, 3)$.

Ans. $y=\frac{\sqrt{3}}{3}x+3$.

5. Find the equation of the straight line forming an angle of 60° with Ox and meeting Oy in the point $(0, -4)$.

Ans. $y=\sqrt{3}\,x-4$.

6. Find the equation of the straight line forming an angle of 120° with Ox and meeting Oy in the point $(0, 5)$.

Ans. $y=-\sqrt{3}\,x+5$.

7. Find the equation of the straight line parallel to Ox that meets Oy in the point $(0, 3)$.

Ans. $y=3$.

8. Find the equation of the straight line parallel to Oy that meets Ox in the point $(-5, 0)$.

Ans. $x=-5$.

9. Find the equation of the straight line passing through the point $(1, -3)$ and forming an angle of arc tan 2 with Ox.

Ans. $2x-y-5=0$.

10. Find the equation of the straight line passing through the point $\left(-1, -\frac{1}{2}\right)$ and forming an angle of arc tan (-2) with Ox.

Ans. $4x+2y+5=0$.

11. Find the equation of the straight line passing through the point $\left(2, \frac{5}{3}\right)$ and forming an angle of 0° with Ox.

Ans. $y=-\frac{5}{3}$.

12. Find the equation of the straight line passing through the point $\left(\frac{1}{3}, \frac{2}{3}\right)$ and parallel to Oy.

Ans. $x=\frac{1}{3}$.

13. Find the slopes of the straight lines:
(*a*) $x-y-5=0$; (*b*) $6x-3y+7=0$; (*c*) $3x+2y-1=0$.

Ans. (*a*) 1 ; (*b*) 2 ; (*c*) $-\frac{3}{2}$.

14. Find the equation of the straight line passing through the points $(-1, -4)$ and $(0, 5)$.

Ans. $9x-y+5=0$.

15. Find the equation of the straight line passing through the points $\left(2, -\frac{1}{2}\right)$ and $\left(-1, \frac{1}{4}\right)$.

Ans. $x+4y=0$.

16. Find the equation of the straight line passing through the points $(2, -1)$ and $(2, 3)$.

Ans. $x=2$.

17. Find the equation of the straight line passing through the point $(5, -1)$ and parallel to the straight line joining the points $(0, 3)$ and $(2, 0)$.

Ans. $3x+2y-13=0$.

18. Find the equation of the straight line meeting Ox in the point $(3, 0)$ and Oy in the point $(0, -4)$.

Ans. $4x-3y-12=0$.

19. Find the points in which the straight line $6x-4y-3=0$ meets the axes.

Ans. $\left(\frac{1}{2}, 0\right)$, $\left(0, -\frac{3}{4}\right)$.

20. The diagonals of a rhombus are 12 and 8 units of length and lie on the axes. Find the equations of the sides of the rhombus.

Ans.
$$\begin{aligned} &2x+3y-12=0, \\ &2x-3y+12=0, \\ &2x+3y+12=0, \\ &2x-3y-12=0, \end{aligned}$$

or

$$\begin{aligned} &3x+2y-12=0, \\ &3x-2y+12=0, \\ &3x+2y+12=0, \\ &3x-2y-12=0. \end{aligned}$$

21. Find the area of the triangle formed by the axes and the straight line $3x+4y-12=0$.

Ans. 6 sq. units.

22. What relationship must hold between the coefficients a and b for the straight line

$$\frac{x}{a}+\frac{y}{b}=1$$

o form with Ox an angle of (a) 45°; (b) 60°; (c) 135°.

Ans. (a) $a=-b$; (b) $a=-\frac{b\sqrt{3}}{3}$; (c) $a=b$.

23. Find the acute angle α between the straight lines $3x-y+6=0$ and $x-y+4=0$.

Ans. $\alpha=\text{arc tan}\,\frac{1}{2}$.

24. Find the acute angle α between the straight lines $2x-y+8=0$ and $2x+5y-4=0$.

Ans. $\alpha=\text{arc tan}\ 12$.

25. Find the acute angle between the straight line $2x-3y+6=0$ and the straight line passing through the points $(4, -5)$ and $(-3, 2)$.

Ans. arc tan 5.

26. Find the angle α between the straight lines passing through the origin and the points that trisect the intercept of the straight line $2x+3y-12=0$ contained between the axes.

Ans. $\alpha=\text{arc tan}\,\frac{9}{13}$.

27. Prove that the straight line $x-y+3=0$ is a bisector of one of the angles between the straight lines

$$4x-3y+11=0 \text{ and } 3x-4y+10=0 .$$

28. Find the equation of the straight line passing through the point $(2, -3)$ and parallel to the straight line $3x-2y+2=0$.

Ans. $3x-2y-12=0$.

29. Find the equation of the straight line passing through the point $\left(-\frac{3}{2}, -2\right)$ and parallel to the straight line $3x-2y+2=0$.

Ans. $6x-4y+1=0$.

30. Find the equation of the straight line through the point $(-1, -1)$ and parallel to the straight line through the points $(-2, 6)$ and $(2, 1)$.

Ans. $5x+4y+9=0$.

31. Find the equation of the straight line through the origin which is perpendicular to the straight line $3x+4y-2=0$.

Ans. $4x-3y=0$.

32. Find the equation of the straight line passing through the point $(2, -3)$ and perpendicular to the straight line $7x-4y+3=0$.

Ans. $4x+7y+13=0$.

33. Find the equation of the perpendicular bisector of the line joining the points $(-5, -1)$ and $(-3, 4)$.

Ans. $4x+10y+1=0$.

34. Find the equation of the perpendicular to the straight line $2x-3y+7=0$ passing through the mid-point of its intercept contained between the axes.

Ans. $36x+24y+35=0$.

35. Find the equation of the straight line through the point $(4, -3)$ that forms an angle of $45°$ with the straight line $3x+4y=0$.

Ans. $x-7y-25=0$.

36. Find the equation of the straight line through the point $(-1, -1)$ that forms an angle of arc tan $\frac{1}{2}$ with the straight line $3x+2y-6=0$.

Ans. $4x+7y+11=0$.

37. Find the equations of the perpendiculars to the straight line $y=3x+1$ erected on its points of intersection with the axes.

Ans. $3x+9y+1=0$; $x+3y-3=0$.

38. Find the straight line through the point $(3, -5)$ that forms an angle with Ox equal to twice the angle that the straight line $x-2y-5=0$ forms with Ox.

Ans. $4x-3y-27=0$.

39. Find the equation of the straight line through the point $\left(\frac{2}{3}, \frac{8}{3}\right)$ and through the point of intersection of the straight lines $3x-5y-11=0$ and $4x+y-7=0$.

Ans. $11x+4y-18=0$.

40. Find the equation of the straight line through the points of intersection of the straight lines

$$2x-y-1=0, \quad x-y+7=0$$

and

$$x-7y-1=0, \quad 2x-5y+1=0.$$

Ans. $23x-14y+26=0$.

41. Find the equation of the straight line passing through the point of intersection of the straight lines $x-3y+2=0$ and $5x+6y-4=0$ and parallel to the straight line $4x+y+7=0$.

Ans. $12x+3y-2=0$.

42. Find the equation of the straight line passing through the point of intersection of the straight lines $3x-y+4=0$, $4x-6y+3=0$ and perpendicular to the straight line $5x+2y+6=0$.

Ans. $4x-10y+1=0$.

43. Find the equations of the medians of the triangle formed by the straight lines $2x-3y+11=0$, $3x+y-11=0$ and $x+4y=0$.

Ans. $5x-2y=0$; $4x+5y-11=0$; $x-7y+11=0$.

44. Find the equation of the straight line passing through the origin and the point of intersection of the medians of the triangle whose sides are given by the equations

$$4x-y+4=0, \quad y=-x+4, \quad x-4y+1=0.$$

Ans. $5x-2y=0$.

45. Find the base of the perpendicular from the point $(-1, 2)$ to the straight line $3x-5y-21=0$.

Ans. $(2, -3)$.

46. Find the equation of the perpendicular to the straight line through the points $M_1(-4, 6)$ and $M_2(4, -1)$ which meets the intercept M_1M_2 in a point whose distance from M_1 is one-third the length of M_1M_2.

Ans. $24x-21y+109=0$.

47. Find the distance from the point $(2, 1)$ to the straight line $3x-y+7=0$.

Ans. $\frac{6\sqrt{10}}{5}$.

48. Find the distance from the point $\left(2, -\frac{3}{2}\right)$ to the straight line $x+2y-4=0$.

Ans. $\sqrt{5}$.

49. Find the distance between the parallel straight lines $2x+3y-8=0$ and $2x+3y-10=0$.

Ans. $\frac{2}{13}\sqrt{13}$.

50. The straight line joining $M_1(-3, 1)$ and $M_2(5, -1)$ forms the base of a triangle. Find the height of the triangle measured from its third vertex $M_3(6, 5)$.

Ans. $\frac{25}{17}\sqrt{17}$.

51. For what value of the coefficient m does the straight line $y=mx+3$ pass through the point of intersection of the straight lines $2x-y+1=0$ and $y=x+5$?

Ans. $\frac{3}{2}$.

52. From the point $(9, 5)$ three perpendiculars are let fall on to the sides of a triangle whose vertices are $(8, 8)$, $(0, 8)$, $(4, 0)$. Show that the bases of the three perpendiculars lie on a straight line.

53. Find the point on the straight line $2x+3y-6=0$ equidistant from $(4, 4)$ and $(6, 1)$.

Ans. $\left(\frac{17}{8}, \frac{7}{12}\right)$.

54. Find the point on the straight line $5x-3y+15=0$ whose distance from Ox is $\frac{2}{3}$ its distance from Oy.

Ans. $\left(-5, -\frac{10}{3}\right)$, $\left(-\frac{15}{7}, \frac{10}{7}\right)$.

55. A point M is at a distance of 8 units from the origin, and the slope of the straight line joining M and the origin is $-\frac{1}{4}$. Find the co-ordinates of M.

Ans. $\left(\pm\frac{32}{17}\sqrt{17}, \pm\frac{8}{17}\sqrt{17}\right)$.

56. A point M is at a distance of 5 units from the point $(1, -2)$, whilst the slope of the straight line passing through M and the point $(0, -8)$ is equal to $\frac{1}{2}$. Find the co-ordinates of M.

Ans. $(4, -6)$; $\left(\frac{12}{5}, -\frac{34}{5}\right)$.

57. Show analytically that the locus of points equidistant from two given points is the perpendicular bisector of the line joining the points.

58. Show analytically that the medians of a triangle meet in a single point.

59. Show analytically that the perpendiculars from the vertices of a triangle to its opposite sides meet in a single point.

60. A beam of light $y=x+3$ is incident on a glass plate of thickness 1 cm (index of refraction$=2$). Assuming that the plate is arranged so that the axis of abscissal lies on the upper surface of the plate and the axis of ordinates is perpendicular to it, find the equation of the beam during its passage through the plate and after leaving on the other side; find also the length of the path of the beam inside the plate.

Ans. $y=\sqrt{7}\,(x+3)$; $\quad y=x+2+\dfrac{1}{\sqrt{7}}$; $\quad \dfrac{2\sqrt{2}}{\sqrt{7}}$.

CHAPTER 3

LOCUS. THE EQUATIONS TO LOCI. CURVES OF THE SECOND ORDER

§ 16. Locus. The equation of a curve specified as a locus. Second-order curves. 1. The method described in Chapter 2 for obtaining the equation of a straight line in accordance with the geometrical conditions that define the line can be extended to curves in general if they are specified as loci.

The set of points on a plane possessing some property that distinguishes them from all other points of the plane is referred to as a locus. For instance, the set of points with the property that they are all situated at the same distance from a given point forms the locus that we term a circle. The set of points equidistant from two given points forms a straight line, which is the perpendicular bisector of the straight line joining the two given points.

Suppose that a curve is given as a locus. As in the case of a straight line, we can find the equation of the curve by writing in co-ordinate form the geometrical equality expressing the condition that the variable point $M(x, y)$ belongs to the locus.

2. We shall consider an example of finding the equation of a curve given as a locus.

Example. To find the locus of a point which moves so that its distance from a given straight line AB is half its distance from a given point P, which does not lie on AB.

Solution. To find the equation of a locus we must first of all decide on the position of the co-ordinate axes. Though the position of the axes can obviously be chosen arbitrarily, the equation of the locus will be simpler if the choice is a good one. There are no rules of guidance in this matter and skill in making a suitable choice for the position of the co-ordinate axes is purely a matter of practice.

We take the given straight line AB as the x-axis whilst the y-axis is taken through our given point P (fig. 19). Since the point P is given we must know its distance from AB, i.e. from the x-axis. Let a denote this distance. Then the point P has the co-ordinates $(0, a)$.

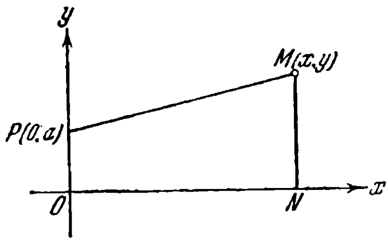


FIG. 19

The condition that the point $M(x, y)$ belongs to the locus is expressed by the equation

$$PM = 2.\ NM\ .$$

To find the equation of the locus we must express the above equation in co-ordinate form, i.e. write the lengths of MP and MN in terms of the co-ordinates of their ends, M, P and N.

We obtain from the expression for the distance between two points [expression (3), Chapter 1]:

$$PM = \sqrt{x^2 + (y-a)^2}\ .$$

The length of NM is equal to the absolute value of the ordinate y of the point M: $NM = |y|$ (the ordinate y of M may be negative, whereas the length of NM is by its nature a positive quantity). Thus

$$\sqrt{x^2 + (y-a)^2} = 2\,|\,y\,|\,.$$

On squaring both sides of this equation and collecting like terms, we obtain for the equation of the locus:

$$x^2 - 3y^2 - 2ay + a^2 = 0\ .$$

This locus is illustrated in fig. 20.

If we took a different position for the co-ordinate axes, so that, for instance, the y-axis no longer passes through the point P, we

should have a non-zero value, say b, for the abscissa of P. We should get in this case:

$$\sqrt{(x-b)^2+(y-a)^2}=2\,|\,y\,|\,,$$

or

$$x^2-3y^2-2bx-2ay+a^2+b^2=0\,,$$

whence it will be seen that a more complicated equation is obtained.

3. *Loci which are expressed by equations of the second degree are called curves of the second order.* Such curves include the circle

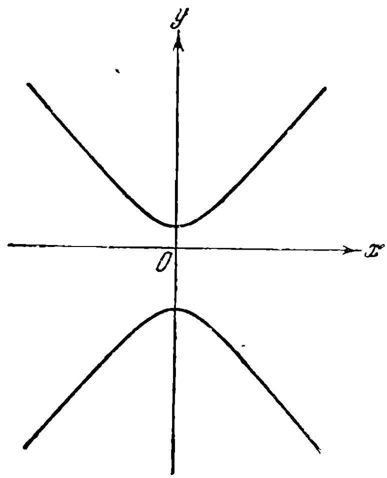


FIG. 20

and the following: the ellipse, hyperbola, and parabola.* We turn now to finding the equations of these curves, after which we study their properties and plot them with the aid of the equations obtained.

§ **17. The circle.** 1. We deduce the equation of the circle of radius r with centre at the point $C\,(a\,,\,b)$. The condition that the variable point $M\,(x\,,\,y)$ lies on the circle is expressed by the equation

* In addition to the curves mentioned there are the so-called degenerate curves of the second order; an investigation of these is bound up with the general study of equations of the second degree and lies outside the scope of the present course.

$$CM = r$$

or

$$CM^2 = r^2$$

(fig. 21). It remains to find the equation of the circle by writing

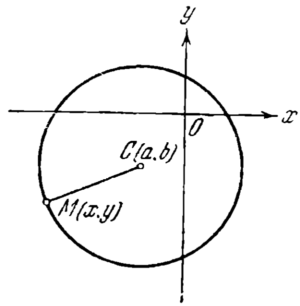


FIG. 21

the above equation in co-ordinate form. We find on using the expression for the distance between two points:

$$(x-a)^2+(y-b)^2=r^2 . \qquad (1)$$

This in fact represents the equation of the circle of radius r with centre at the point $C(a, b)$.

If the centre of the circle is located at the origin, we have $a=b=0$, and equation (1) takes the simpler form

$$x^2+y^2=r^2 . \qquad (1^*)$$

Example. Write the equation of the circle of radius $\frac{2}{3}$ with centre at the point $\left(\frac{1}{2}, -\frac{1}{3}\right)$.

Solution. We obtain by equation (1)

$$\left(x-\frac{1}{2}\right)^2+\left(y+\frac{1}{3}\right)^2=\frac{4}{9} .$$

If we now remove the brackets and simplify, we arrive at the following (equivalent) equation:

$$x^2+y^2-x+\frac{2}{3}y-\frac{1}{12}=0$$

or, writing without fractions:

$$12x^2+12y^2-12x+8y-1=0\ .$$

2. By proceeding as in the previous section we can reduce any equation of a circle of the form (1) to the equivalent form

$$Ax^2+Ay^2+Bx+Cy+D=0\ ;$$

this latter equation is known as *the general equation of a circle.*

Thus *every circle can be expressed by means of an equation of the form* (2).

The most general form of equation of the second degree in two variables x and y is

$$Mx^2+Nxy+Py^2+Qx+Ry+S=0\ .$$

On comparing equation (2) for a circle with this latter equation, we arrive at the conclusion that the general equation of a circle is an equation of the second degree in two variables x and y with the following special features: (*i*) the coefficients of the squares of the current co-ordinates are equal ($M=P$); (*ii*) the term containing the product xy of the current co-ordinates is missing ($N=0$).

3. We now take the converse problem and consider what geometric curves can correspond to equations of the form (2).

Suppose, for instance, we are given the equation

$$9x^2+9y^2+12x-18y-23=0\ . \tag{3}$$

This equation is of the form (2); the coefficients of the squares of the current co-ordinates are equal and the term containing the product xy is absent; here

$$A=9\ , \quad B=12\ , \quad C=-18 \text{ and } D=-23\ .$$

We reduce equation (3) to form (1). We do this by dividing all the terms of equation (3) by 9 and taking the free term over to the right-hand side; we get the equivalent equation

$$x^2+\frac{4}{3}x+y^2-2y=\frac{23}{9}. \tag{3*}$$

We now consider what numbers must be added to the expressions

$$x^2+\frac{4}{3}x \quad \text{and} \quad y^2-2y$$

in order to form the sum of two complete squares on the left-hand side of equation (3*), as in equation (1).

Observing that $\frac{4}{3}x=2\cdot\frac{2}{3}x$, we shall regard the expression $x^2+ +2\cdot\frac{2}{3}x$ as the sum of the square of the "first term" and twice the product of the "first term" with the "second". Since the "first term" is x (its square is x^2), we conclude from the product $2\cdot\frac{2}{3}x$ that the "second term" is the number $\frac{2}{3}$; this means that we form a complete square by adding to $x^2+2\cdot\frac{2}{3}x$ the number $\frac{4}{9}$:

$$x^2+2\cdot\frac{2}{3}x+\frac{4}{9}=\left(x+\frac{2}{3}\right)^2.$$

Similarly, by adding 1 to the difference $y^2-2y=y^2-2.1.y$, we form the second complete square $(y-1)^2$.

Thus we have to add $\frac{4}{9}$ and 1 to the left-hand side of equation (3*) to form the sum of complete squares $\left(x+\frac{2}{3}\right)^2+(y-1)^2$, so that we must add the same sum $\frac{4}{9}+1=\frac{13}{9}$ to the right-hand side in order to obtain an equation equivalent to (3*).

Hence we obtain the equation

$$\left(x^2+\frac{4}{3}x+\frac{4}{9}\right)+(y^2-2y+1)=\frac{23}{9}+\frac{13}{9}$$

or

$$\left(x+\frac{2}{3}\right)^2+(y-1)^2=4\,. \qquad (3^{**})$$

If we take the circle of radius 2 with centre at the point $\left(-\frac{2}{3}, 1\right)$, expression (1) shows us that it will have the equation (3**); hence equation (3**) represents a circle; and since equation (3**) is equivalent to (3), this latter equation must also represent a circle.

We now consider another equation of type (2):

$$x^2+y^2-2x+4y+5=0\,.$$

If we proceed as in the previous example, we arrive at the following equivalent equation:

$$(x-1)^2+(y+2)^2=0\,.$$

A unique pair of co-ordinates $(1\,,\,-2)$ satisfies this equation, since the sum of two non-negative quantities can only vanish if each quantity vanishes separately. The given equation thus represents a point, and not a circle.

Finally, if we take the equation

$$x^2+y^2-2x+4y+7=0\,,$$

it reduces to the following equivalent equation:

$$(x-1)^2+(y+2)^2=-2\,. \qquad (3^{***})$$

There are no real values for which the sum of two non-negative magnitudes can become negative; hence equation (3***) is not satisfied for any real values of the co-ordinates $(x\,,y)$. This means that there is no geometric shape on the xOy plane corresponding to equation (3***), or, consequently, to our initial equation.

When any equation (2) is transformed to the equivalent equation (1) the right-hand side of the latter can consist of either a positive number, or zero, or a negative number. In the first case, as we have seen in the examples, equation (2) represents a circle; in the second case it represents a point, and in the third, no geometric shape of any kind.

We thus deduce the following:

Equation (2) *represents either a circle, or a point, or no geometric shape whatever.*

We can proceed as in the above examples in order to find out which of these three cases applies for any given equation of type (2), i.e. we reduce the equation of type (2) to its equivalent of type (1).

§ 18. The ellipse. The locus of a point which moves so that the sum of its distances from two given points called the foci is constant is called an ellipse. The constant in question must be greater than the distance between the foci.

We find the equation of the ellipse, i.e. we find the equation which the co-ordinates of any point of the ellipse must satisfy.

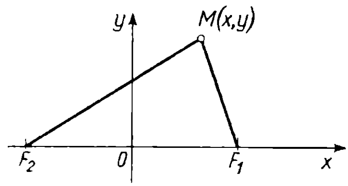


FIG. 22

Let F_1, F_2 denote the foci. We take the x-axis along the straight line joining F_1 and F_2, and the origin at the mid-point of F_1F_2 (fig. 22).

Let $2a$ denote the constant sum of the distances of any point of the ellipse from the foci. The condition that a point $M(x, y)$ lies on the ellipse is now expressed by the equation

$$F_1M + F_2M = 2a\,. \tag{4}$$

To write this equation in co-ordinate form we need to know the co-ordinates of the foci F_1, F_2; let the distance between the foci be equal to $2c$. Then the co-ordinates of F_1 will be c and 0, whilst those of F_2 will be $-c$ and 0. We now use the expression for the distance between two points, and find that

$$F_1M = \sqrt{(x-c)^2 + y^2} \quad \text{and} \quad F_2M = \sqrt{(x+c)^2 + y^2}\,,$$

On substituting these expressions in equation (4), we get for the equation of the ellipse:

$$\sqrt{(x-c)^2+y^2}+\sqrt{(x+c)^2+y^2}=2a\ . \tag{5}$$

Equation (5) can be reduced to a simpler form by getting rid of the radicals. We first take the second radical over to the right-hand side:

$$\sqrt{(x-c)^2+y^2}=2a-\sqrt{(x+c)^2+y^2}\ . \tag{5*}$$

We square both sides of equation (5*):

$$x^2-2cx+c^2+y^2$$
$$=4a^2-4a\sqrt{(x+c)^2+y^2}+x^2+2cx+c^2+y^2\ .$$

On re-arranging and cancelling, and dividing both sides by 4, we arrive at

$$a\sqrt{(x+c)^2+y^2}=a^2+cx\ .$$

We again square both sides:

$$a^2(x^2+2cx+c^2+y^2)=a^4+2a^2cx+c^2x^2\ ,$$

whence we obtain on re-arranging and simplifying:

$$a^2x^2-c^2x^2+a^2y^2=a^4-a^2c^2\ ,$$

or

$$(a^2-c^2)x^2+a^2y^2=a^2(a^2-c^2)\ . \tag{5**}$$

This equation has a simpler form than (5).

This last equation is usually written in yet another form.

Since $2a>2c$ by hypothesis, i.e. $a>c$, the difference a^2-c^2 is positive and can therefore be denoted by $b^2(b=\sqrt{a^2-c^2})$:

$$a^2-c^2=b^2\ . \tag{6}$$

On replacing a^2-c^2 by b^2 in equation (5**) for the ellipse, we get

$$b^2x^2+a^2y^2=a^2b^2\ , \tag{7}$$

or, after dividing through by a^2b^2:

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1. \tag{8}$$

The equation obtained is known as *the canonical equation of the ellipse.** If the co-ordinate axes had been located differently, a more complicated equation would have been obtained.

§ 19. The shape of the ellipse. We find the shape of the ellipse by first solving equation (8) with respect to the ordinate y:

$$\frac{y^2}{b^2}=1-\frac{x^2}{a^2} \quad \text{or} \quad y^2=\frac{b^2}{a^2}(a^2-x^2),$$

whence

$$y=\pm\frac{b}{a}\sqrt{a^2-x^2}. \tag{9}$$

It is clear from this equation that, for each value of the abscissa x with $x^2<a^2$, there are two corresponding real values of the ordinate y, equal in absolute value but opposite in sign; this shows that the x-axis is an axis of symmetry of the ellipse.

Similarly, if we solve the equation of the ellipse for the abscissa x, i.e. write the equation

$$x=\pm\frac{a}{b}\sqrt{b^2-y^2}, \tag{9*}$$

we can conclude that the y-axis is also an axis of symmetry of the ellipse.

It follows from equation (9) that the ordinate y only takes real values for values of x satisfying the relationship $x^2\leqq a^2$, i.e. when the absolute value of x does not exceed a. Similarly, if follows from equation (9*) that the abscissa x is only real when $y^2\leqq b^2$, i.e. when the absolute value of y does not exceed b.

* Cf. the first foot-note on p. 44 regarding the equivalence of equation (8) to the initial equation (5).

The points, whose abscissae have absolute values not exceeding a, lie in an infinite strip bounded by the straight lines PQ and SR, parallel to the y-axis and at equal distances a from it, the one on the right and the other on the left (fig. 23). The points, whose

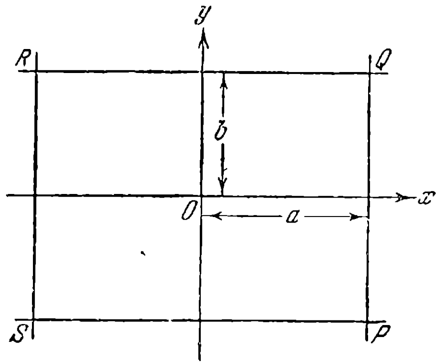


FIG. 23

ordinates have absolute values not exceeding b, lie in a strip bounded by the straight lines RQ and SP, parallel to the x-axis and at equal distances b from it, the one above and the other below. The points whose abscissae and ordinates simultaneously satisfy the conditions in question lie in the area common to these strips, i.e. in the rectangle $PQRS$ (fig. 23).

It is clear from equation (9*) that the greatest absolute value for the abscissa of a point of the ellipse is obtained when $y=0$. If $y=0$, $x= \pm a$. The ellipse thus cuts the x-axis in the points $(a, 0)$ and $(-a, 0)$. We see from equation (9) that the greatest absolute value of the ordinate of a point of the ellipse corresponds to $x=0$. When $x=0$, $y= \pm b$. Hence the ellipse cuts the y-axis in the points $(0, b)$ and $(0, -b)$.

We see from equation (9) that, as the abscissa x of a point of the ellipse increases fiom 0 to a, the ordinate y of the point decreases from b to 0.

Bearing in mind the above remarks about the symmetry of the ellipse with respect to the axes Ox and Oy, we arrive at the conclusion that the ellipse has a shape similar to that shown in fig. 24. (The points M_1, M_2, M_3 and M_4 are shown on the figure in order to

bring out more clearly the symmetry of the ellipse with respect to the co-ordinate axes).

We call $A'A$ the *major axis* of the ellipse. The length of the major axis is $2a$. We call $B'B$ the *minor axis* of the ellipse; its length

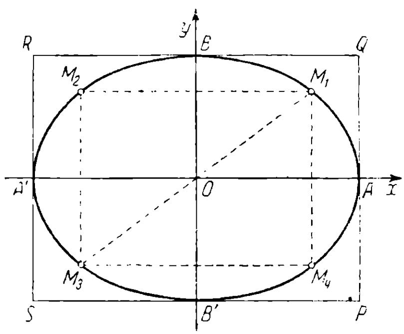


FIG. 24

is $2b$. The points A, A', B, B' are known as the *vertices* of the ellipse, whilst its *centre* is defined as the point of intersection 0 of its axes.

Note. With $b>a$ equation (8) clearly becomes the equation of the ellipse whose foci lie on the Oy axis at distances of $\sqrt{b^2-a^2}$ from the origin. In this case $B'B=2b$ is the major axis and $A'A=2a$ is the minor axis.

§ 20. Eccentricity of the ellipse. Connexion between the ellipse and the circle. The ratio of the distance $2c$ between the foci of an ellipse and the length $2a$ of the major axis, i.e. $\frac{2c}{2a}=\frac{c}{a}$, is called the *eccentricity* and is usually denoted by the letter ε. We have by equation (6)

$$c=\sqrt{a^2-b^2}.$$

Hence

$$\varepsilon=\frac{c}{a}=\frac{\sqrt{a^2-b^2}}{a}. \tag{10}$$

It follows from the definition that the eccentricity is always less than unity.

We have from equation (10)

$$\varepsilon=\sqrt{\frac{a^2-b^2}{a^2}} \quad \text{or} \quad \varepsilon=\sqrt{1-\left(\frac{b}{a}\right)^2}.$$

It follows from the last relationship that the closer the eccentricity is to 1, the smaller is the ratio $\frac{b}{a}$; and the smaller the ratio of the lengths of the semi-axes, the more elongated the ellipse. Hence the eccentricity characterizes the shape of the ellipse.

We consider the particular case of the ellipse when $a=b$. Equation (8) now has the form

$$\frac{x^2}{a^2}+\frac{y^2}{a^2}=1, \quad \text{or} \quad x^2+y^2=a^2.$$

But this is the equation of the circle of radius a with centre at the origin. We conclude from this that the circle is the particular case of the ellipse when the axes of the latter are equal.

It follows from (10) that, with $a=b$, the eccentricity

$$\varepsilon=\frac{\sqrt{a^2-b^2}}{a}=0.$$

Hence the circle is the ellipse whose eccentricity is zero.

§ **21. The hyperbola. The hyperbola is defined as the locus of a point which moves so that the difference between its distances from two fixed points (called the foci) has a constant absolute value.** The constant value in question must be non-zero and less than the distance between the foci.

Let F_1 and F_2 be the foci (fig. 25). We take the same position of the co-ordinate axes relative to the foci as when finding the equation of the ellipse. We write $2c$ for the distance between the foci F_1 and F_2, and $2a$ for the difference between the distances of any point of the hyperbola from the foci, noting that $2a<2c$ by the definition of the hyperbola, i.e. $a<c$. The condition that a point $M(x, y)$ lies on the hyperbola is given by

$$F_2 M - F_1 M = \pm 2a \,^*.$$

To obtain the equation of the hyperbola, it remains to write this expression in co-ordinate form. This is easily done by writing the

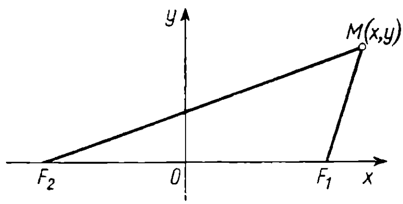


FIG. 25

lengths of F_2M and F_1M in terms of the co-ordinates of their ends. We obtain

$$\sqrt{(x+c)^2+y^2} - \sqrt{(x-c)^2+y^2} = \pm 2a \,.$$

Instead of one equation we have obtained two, each of which expresses the condition that the point $M(x\,,y)$ lies on the hyperbola; this points to the fact the hyperbola consists of two parts, or as we usually say, of two branches.

We re-write the equations by taking the second radical over to the right-hand side:

$$\sqrt{(x+c)^2+y^2} = \pm 2a + \sqrt{(x-c)^2+y^2} \,.$$

We obtain by squaring both sides

$$x^2+2cx+c^2+y^2 = 4a^2 \pm 4a\sqrt{(x-c)^2+y^2} + x^2 - 2cx + c^2 + y^2$$

or, on simplifying:

$$\pm a\sqrt{(x-c)^2+y^2} = a^2 - cx \,.$$

On again squaring both sides, we arrive at the equation

$$(a^2-c^2)\,x^2 + a^2\,y^2 = a^2\,(a^2-c^2) \,. \tag{11}$$

---

* $F_2M - F_1M$ is positive if the point M is taken so that $F_2M > F_1M$ (as in fig. 25), and is negative when $F_2M < F_1M$.

Equation (11) has the same form as (5**) for the ellipse before introducing the quantity b. But there is the difference that now $a<c$, so that a^2-c^2 is a negative quantity, which can be denoted by $-b^2$.

We obtain on replacing a^2-c^2 by $-b^2$ in equation (11):

$$-b^2x^2+a^2y^2=-a^2b^2.$$

On dividing through by $-a^2b^2$, this reduces to the form

$$\frac{x^2}{a^2}-\frac{y^2}{b^2}=1. \qquad (12)$$

This last is called the *canonical equation of the hyperbola.**

From $a^2-c^2=-b^2$ it follows that

$$c=\sqrt{a^2+b^2}.$$

The ratio $\frac{c}{a}$ is called the *eccentricity of the hyperbola* and is usually denoted by the letter ε:

$$\varepsilon=\frac{c}{a}=\frac{\sqrt{a^2+b^2}}{a}.$$

It follows from the definition that the eccentricity of a hyperbola is greater than unity.

We find the shape of the hyperbola by solving its equation (12) for the ordinate y:

$$\frac{y^2}{b^2}=\frac{x^2}{a^2}-1, \qquad y^2=\frac{b^2}{a^2}(x^2-a^2)$$

and

$$y=\pm\frac{b}{a}\sqrt{x^2-a^2}. \qquad (12^*)$$

---

* The same remark can be made regarding the working leading up to this equation as was made when we deduced the equation of the straight line (cf. the first foot-note on p. 44).

We see from this that, for each value of the abscissa x with $x^2 > a^2$ there are two corresponding real values of the ordinate y, equal in absolute value and opposite in sign. This shows that Ox is an axis of symmetry of the hyperbola.

Similarly, we conclude from the equation for the hyperbola solved for the abscissa x, i.e. from the equation

$$x = \pm \frac{a}{b} \sqrt{y^2 + b^2}, \tag{12**}$$

that Oy is also an axis of symmetry of the hyperbola.

As equation (12*) shows, y only takes real values for values of x satisfying $x^2 \geqq a^2$. If the absolute value of the abscissa is less than a, y becomes an imaginary quantity. Hence it follows that there are no points of the hyperbola in the strip bounded by the straight lines RS and PQ, parallel to the axis of ordinates and at distances a from it on the left and the right. (See fig. 26.)

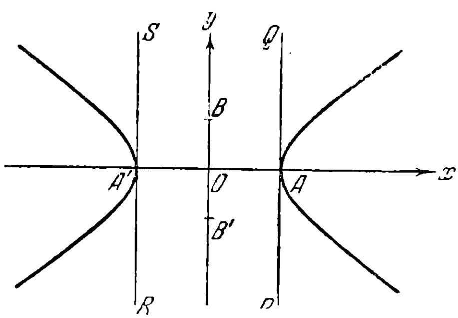


FIG. 26

For $x = \pm a$, the ordinate $y = 0$. This means that the hyperbola intersects Ox in the points $A(a, 0)$ and $A'(-a, 0)$; these points have the least absolute value of abscissa of any points on the hyperbola.

It follows from equation (12*) that the absolute value of the ordinate y of a point of the hyperbola increases as the absolute value of the abscissa x increases. This indicates that the branches of the hyperbola extend indefinitely to the right and left from the straight lines RS and PQ and upwards and downwards from the axis of abscissae.

We conclude from all that has been said that the hyperbola has the shape shown in fig. 26.

We call AA' the *transverse axis* of the hyperbola, whilst A and A' are known as its vertices. BB', of length $2b$, is known as the *conjugate axis* of the hyperbola.

The point O, equidistant from and lying on the straight line joining the vertices, is called the *centre* of the hyperbola.

§ 22. Asymptotes of the hyperbola. We consider more closely the nature of the growth of the absolute value of the ordinate y of a point of the hyperbola when the absolute value of its abscissa x increases indefinitely. In view of the symmetry of the hyperbola with respect to the co-ordinate axes it is sufficient to consider this question for the part of the right-hand branch lying in the first quadrant, i.e. for positive values of x and y.

We write the equation of this branch in the form solved for y:

$$y = \frac{b}{a}\sqrt{x^2 - a^2}\ ;$$

whence

$$y = \frac{b}{a}\,x\sqrt{1 - \frac{a^2}{x^2}}\,. \tag{12***}$$

The fraction $\dfrac{a^2}{x^2}$ is very small for large values of x, so that the value of the radical $(1 - a^2/x^2)$ is close to 1. This fact suggests that we compare the variation in the ordinate y of a point of the hyperbola with the variation in the ordinate of a point of the straight line $y = \dfrac{b}{a}x$, the equation of which is obtained from (12***) by replacing the radical by unity.

To distinguish between the ordinate y of a point of the hyperbola from the ordinate of a point on the straight line (when both the points in question have the same abscissa x), we shall write Y for the ordinate of the point of the straight line. We now investigate

the difference $Y-y$; since $Y=\frac{b}{a}x$, and $y=\frac{b}{a}\sqrt{x^2-a^2}$, we have

$$Y-y=\frac{b}{a}x-\frac{b}{a}\sqrt{x^2-a^2}=\frac{b}{a}(x-\sqrt{x^2-a^2}).$$

As x increases, $Y-y$ varies, though it is difficult to tell from the expression obtained whether it increases or decreases. We therefore transform the right-hand side by multiplying top and bottom by $x+(x^2-a^2)$, which gives us:

$$Y-y=\frac{b(x-\sqrt{x^2-a^2})(x+\sqrt{x^2-a^2})}{a(x+\sqrt{x^2-a^2})}$$

or, after simplifying:

$$Y-y=\frac{ab}{x+\sqrt{x^2-a^2}}.$$

We can now see from the expression on the right-hand side that $Y-y$ diminishes indefinitely as the abscissa x increases, for the numerator of the fraction remains constant whilst the denominator increases indefinitely. This means that a point of the hyperbola moving along the branch in question approaches indefinitely the straight line $Y=+\frac{b}{a}x$ without ever actually reaching it.

It follows from the symmetry of the branches that, when a point of the hyperbola moves along the part of the right-hand branch that lies in the fourth quadrant, it will approach indefinitely the straight line $Y=-\frac{b}{a}x$; also, the disposition of the left-hand branch will be similar to that of the right-hand branch as regards the straight lines $Y=\pm\frac{b}{a}x$ (fig. 27).

If a point of a curve having an infinite branch approaches indefinitely a straight line on moving to an infinite distance along this branch, the straight line is called an asymptote to the curve.

We have thus established that the hyperbola has two asymptotes, defined by the equations

$$y = \pm\frac{b}{a}x\,. \tag{13}$$

It is shown in comprehensive courses of analytic geometry that the hyperbola has no other asymptotes.

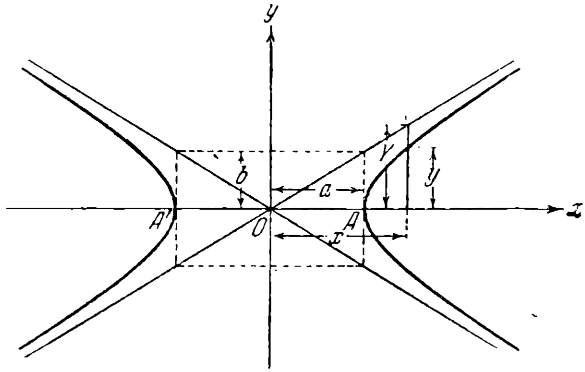


FIG. 27

It follows from the inequality

$$x^2 > x^2 - a^2$$

that the absolute value of x is greater than the absolute value of $\sqrt{(x^2 - a^2)}$. Hence the absolute value of $\frac{b}{a}x$ is greater than that of $\frac{b}{a}\sqrt{(x^2 - a^2)}$, i.e. the absolute value of the ordinate Y on the asymptote is greater than the absolute value of the ordinate y on the hyperbola.

This shows that points of the hyperbola lie inside the vertical angles formed by the asymptotes (fig. 27).

We construct a rectangle on the axes $2a$ and $2b$ of the hyperbola and draw straight lines through pairs of opposite vertices of the rectangle (fig. 27). These straight lines pass through the origin and are diagonals of the rectangle, so that the slope of one is $\frac{b}{a}$ and

of the other $-\dfrac{b}{a}$. Hence they are represented by the equations

$$y=\pm\frac{b}{a}x,$$

which are the same as equations (13) for the asymptotes, i.e. the lines drawn are in fact the asymptotes. This gives us the simplest construction for the asymptotes of a hyperbola.

Example. To find the eccentricity, co-ordinates of the foci, and equations of the asymptotes of the hyperbola

$$\frac{x^2}{25}-\frac{y^2}{4}=1.$$

Solution. We have from the expression deduced in § 21:

$$\varepsilon=\frac{\sqrt{a^2+b^2}}{a}.$$

For the hyperbola in our example, $a=5$, $b=2$; hence

$$\varepsilon=\frac{\sqrt{29}}{5}.$$

The eccentricity is the ratio of the distance of either focus from the origin to the length of the transverse semi-axis (cf. § 21), i.e.

$$\varepsilon=\frac{OF_1}{OA}=\frac{OF_1}{a}$$

(cf. fig. 25). Consequently, $OF_1=a\varepsilon=5\cdot\frac{1}{5}\sqrt{29}=\sqrt{29}$. The foc thus have the co-ordinates $(\sqrt{29}, 0)$ and $(-\sqrt{29}, 0)$.

The equations of the asymptotes are obtained from (13):

$$y=\pm\frac{2}{5}x,$$

or

$$2x\pm 5y=0.$$

§ **23. Rectangular hyperbola.** If $a=b$, the hyperbola is called *rectangular* (or *equilateral*). Its equation has the form

$$\frac{x^2}{a^2}-\frac{y^2}{a^2}=1 .$$

or

$$x^2-y^2=a^2 .$$

From (13), the equations of the asymptotes of a rectangular hyperbola are

$$y=x$$

and

$$y=-x .$$

Thus the asymptotes of a rectangular hyperbola are straight lines, one of which is inclined at 45° to the Ox axis, and the other

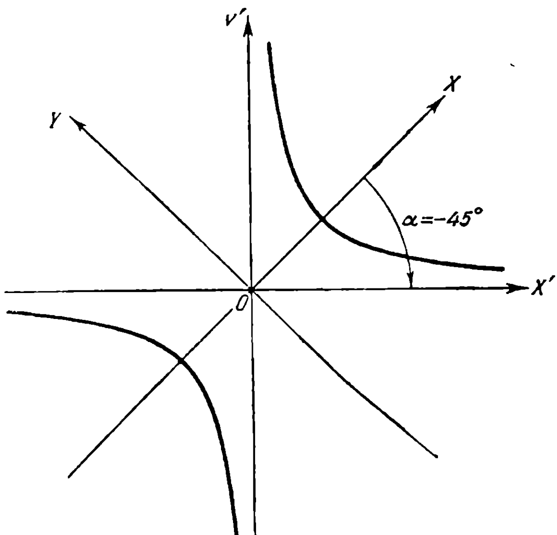


FIG. 28

at 135°. The angle between the asymptotes of a rectangular hyperbola is 90°.

We rotate the co-ordinate axes through the angle $\alpha=-45°$ so that the y-axis now coincides with the asymptote $y=x$, and the x-axis with the asymptote $y=-x$. The asymptotes represent the

new co-ordinate axes. If we draw them in the usual way (with the axis of abscissae horizontal), we obtain fig. 28. Here, Ox' and Oy' denote the new co-ordinate axes.

What does the equation of the rectangular hyperbola

$$x^2 - y^2 = a^2 \tag{14}$$

become with respect to the new co-ordinate axes? To answer this question, we first consider how to express the "old" co-ordinate (x, y) of an arbitrary point M of the plane in terms of the "new" co-ordinates (x', y') of the same point when the old axes are rotated by the angle $\alpha = -45°$.

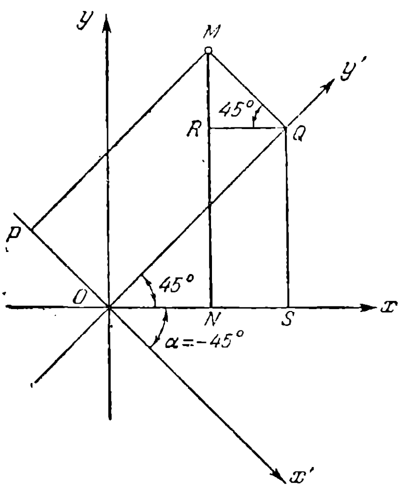


FIG. 29

We have from fig. 29, $x = ON$, $y = NM$, $x' = -OP$ (x' is negative for the position of the point M shown in the figure), $y' = OQ$. Hence:

$$x = ON = OS - NS = OQ \cdot \cos 45° - RQ$$

$$= OQ \cdot \cos 45° - QM \cdot \cos 45° = OQ \cdot \cos 45° + (-OP) \cdot \cos 45°,$$

or

$$x = \frac{\sqrt{2}}{2}(y' + x'). \tag{15}$$

Furthermore

$$y=NM=NR+RM=SQ+RM=OQ\cdot\sin 45^\circ+$$
$$+QM\cdot\sin 45^\circ=OQ\cdot\sin 45^\circ-(-OP)\cdot\sin 45^\circ,$$

or
$$y=\frac{\sqrt{2}}{2}(y'-x'). \qquad (15^*)$$

Expressions (15) and (15*) give the old co-ordinates (x, y) of an arbitrary point of the plane in terms of the co-ordinates (x', y') of the same point with respect to the new co-ordinate system $x'Oy'$. The same relationships will thus apply for points of a hyperbola. In other words, we find the equation of our hyperbola with respect to the new co-ordinate system $x'Oy'$ by replacing x and y in equation (14) by expressions (15) and (15*) respectively. We obtain

$$\left[\frac{\sqrt{2}}{2}(y'+x')\right]^2-\left[\frac{\sqrt{2}}{2}(y'-x')\right]^2=a^2\cdot$$

or, after simplifying:

$$y'x'=\frac{a^2}{2}. \qquad (16)$$

The rectangular hyperbola of equation (14) is thus given by equation (16) when referred to new co-ordinate axes coinciding with its asymptotes; or in other words, equation (16) is *the equation of a rectangular hyperbola referred to its asymptotes.*

If we set $a^2=2m^2$, equation (16) can be written

$$y'=\frac{m^2}{x'},$$

or, on interchanging the roles of "old" and "new" co-ordinate axes:

$$y=\frac{m^2}{x}.$$

We see from this that the graph of inverse proportionality is a rectangular hyperbola referred to its asymptotes.

§ **24. The parabola.** 1. **The locus of a point which moves so that its distance from a given point (the focus) is equal to its distance from a given straigth line (the directrix) is called a parabola** (the focus is assumed not to lie on the directrix).

We take the x-axis perpendicular to the directrix RS and passing through the focus F (fig. 30). We take the origin half way between the focus and directrix, at the mid-point of PF. If we write $PF=p$, the co-ordinates of the focus will be $\left(\frac{1}{2}p, 0\right)$.

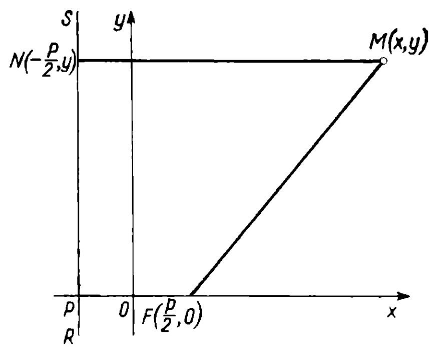


FIG. 30

By definition, the condition that a point $M(x, y)$ lies on the parabola is given by

$$FM=NM.$$

On writing this equality in co-ordinate form we find the equation of the parabola.

The point N has the co-ordinates $\left(-\frac{1}{2}p, y\right)$; hence, using the formula for the distance between two points, the above equality can be written as

$$\sqrt{\left(x-\frac{p}{2}\right)^2+y^2}=\sqrt{\left(x+\frac{p}{2}\right)^2},$$

or, on squaring both sides:

$$x^2-px+\frac{p^2}{4}+y^2=x^2+px+\frac{p^2}{4},$$

whence we get, after cancelling and re-arranging:.

$$y^2 = 2px\,. \tag{17}$$

This is known as the canonical equation of the parabola *.

Equation (17) shows that x can take only non-negative values, since y becomes imaginary for $x<0$. When $x=0$, the ordinate $y=0$. The parabola thus passes through the origin. On re-writing (17) as

$$y = \pm\sqrt{2px}\,,$$

we conclude that Ox is an axis of symmetry of the parabola.

As x increases in absolute value, y also increases. From what has been said, the parabola must have the form shown in fig. 31.

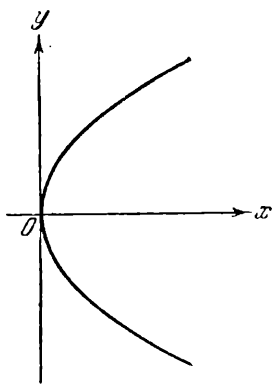


FIG. 31

The point O where the parabola intersects its axis of symmetry is called the *vertex* of the parabola. It is clear from the foregoing that the vertex lies half way between the focus and the directrix.

The quantity p, equal to the distance between the focus and directrix, is known as the parameter or semi-latus rectum of the parabola.

2. If the focus of the parabola is taken to the left of the directrix, its co-ordinates become $(-\frac{1}{2}p\,,0)$ and the equation of the parabola becomes

$$y^2 = -2px\,.$$

* Cf. foot-note on p. 44.

The parabola in this case is situated to the left of the y-axis. We suggest that the reader find the equation of the curve independently for this position of the focus with respect to the directrix.

We suggest as a further exercise for the reader the proof that, if the directrix is horizontal and the focus is situated above the directrix, whilst the y-axis passes through the focus perpendicularly to the directrix, the equation of the parabola takes the form

$$x^2 = 2py .$$

The curve will be symmetrical with respect to Oy. Its branches will be directed upwards. If the focus is below the directrix, the branches will be directed downwards, and the equation of the parabola becomes

$$x^2 = -2py .$$

Example. A stone thrown at an acute angle to the horizon describes the arc of a parabola and falls at a distance of 20 m from the starting-point. The greatest height reached by the stone is 10 m. Find the parameter of the parabolic trajectory.

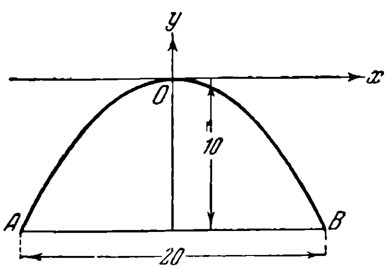


FIG. 32

Solution. On choosing the co-ordinate axes as shown in fig. 32, we see that the trajectory is given by the equation

$$x^2 = -2py .$$

The point B at which the stone falls to earth has the co-ordinates $(10, -10)$. Since this point lies on the parabola, its co-

ordinates (10 , − 10) must satisfy the equation written above. Hence $100 = -2p.(-10)$, or $p=5$.

§ 25. **The parabola** $y = ax^2 + bx + c$. 1. We now let the vertex of the parabola be at any point $O'(\alpha, \beta)$, whilst its axis of symmetry is parallel to Oy (fig. 33).

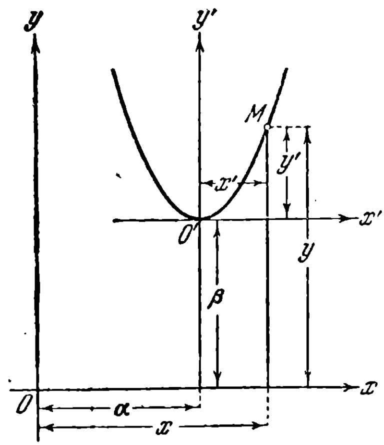


FIG. 33

We construct the auxiliary co-ordinate system $x'O'y'$ for which the origin is at the vertex and the axes are directed parallel to Ox and Oy. Axis $O'y'$ now coincides with the axis of symmetry of the parabola. From the previous section, the equation of the parabola with respect to the new co-ordinate system is

$$x'^2 = 2py' .$$

Our problem amounts to expressing this equation with respect to the original co-ordinate system xOy. We have from fig. 33: $x' = x - \alpha$, $y' = y - \beta$. We obtain, on substituting these values for the current co-ordinates in the above equation:

$$(x-\alpha)^2 = 2p\,(y-\beta) . \tag{18}$$

A parabola has an equation of this form when its vertex is at the point (α, β) and its axis is parallel to Oy.

If the parabola is inverted, with its branches in the negative direction of Oy, its equation will evidently be

$$(x-\alpha)^2=-2p\,(y-\beta)\,. \tag{18*}$$

2. We now show that the curve represented by the equation

$$y=ax^2+bx+c\,, \tag{19}$$

is a parabola. We do this by reducing equation (19) to the form (18).

We divide both sides of (19) by a and take the free term over to the left-hand side; this gives us

$$\frac{y}{a}-\frac{c}{a}=x^2+\frac{b}{a}\,x\,.$$

We now complete the square on the right-hand side by adding $\dfrac{b^2}{4a^2}$ to both sides, whence we have

$$\frac{y}{a}-\frac{c}{a}+\frac{b^2}{4a^2}=x^2+\frac{b}{a}\,x+\frac{b^2}{4a^2}\,,$$

or

$$\frac{1}{a}\left(y-\frac{4ac-b^2}{4a}\right)=\left(x+\frac{b}{2a}\right)^2.$$

On writing

$$\alpha=-\frac{b}{2a}\,,\qquad \beta=\frac{4ac-b^2}{4a}\,,\qquad 2p=\left|\frac{1}{a}\right|\,, \tag{20}$$

we obtain an equation of the form (18) or (18*), i.e.

$$\pm 2p\,(y-\beta)=(x-\alpha)^2.$$

We now consider the parabola for which the co-ordinates $(\alpha\,,\,\beta)$ of the vertex and the parameter p are defined by relationships (20). As we already know, the equation of this parabola will be one of the following:

$$(x-\alpha)^2=\pm 2p\,(y-\beta)\,.$$

These equations are precisely the same as those to which we reduced

equation (19). In other words, equation (19) in fact represents a parabola, whose axis is parallel to Oy.

Example 1. Draw the parabola $y=x^2-4x-5$.

Solution. We find the co-ordinates of the vertex by first adding 9 to both sides of the equation, so that it becomes

$$y+9=(x-2)^2 .$$

The vertex (fig. 34) thus lies at the point $(2, -9)$. To construct the parabola, we find its points of intersection with the co-ordinate

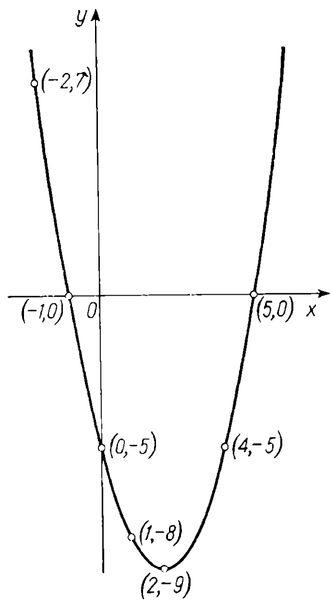


FIG. 34

axes. Setting $x=0$, we get $y=-5$. The curve thus cuts Oy at $(0, -5)$. On setting $y=0$ and solving the quadratic equation

$$x^2-4x-5=0 ,$$

the points of intersection of the parabola with Ox are found to be $(5, 0)$ and $(-1, 0)$. We find a few more points of the curve in order to draw it more accurately. Setting say $x=1$, we get $y=-8$; setting

$x=4$, we get $y=-5$; setting $x=-2$, we get $y=7$, and so on. On plotting these points and joining them with a smooth line, we get the graph shown in fig. 34.

Note. We could have found the co-ordinates of the vertex by using the general expressions derived above, but it is better to proceed as in the example.

Example 2. Find the equation of the parabola passing through the points $(1, 1)$, $(2, 3)$, $(0, 0)$, given that its axis of symmetry is parallel to Oy.

Solution. We write the equation of the parabola in the form

$$y=ax^2+bx+c\,.$$

The co-ordinates of the given points must satisfy the equation of the parabola. On substituting them for the current co-ordinates, we arrive at the system of equations

$$1=a+b+c\,, \quad 3=4a+2b+c\,, \quad 0=c\,,$$

whence we find the coefficients a, b and c:

$$a=\frac{1}{2}\,, \quad b=\frac{1}{2}\,, \quad c=0\,.$$

The required equation is therefore

$$2y=x^2+x\,.$$

§ 26. Curves of the second order as conic sections. Second-order curves can be obtained as the curves of intersection of the surface of a circular cone by planes in different directions. The theory of second-order curves, which is now studied by the methods of analytic geometry, was worked out in detail by the early Greek mathematicians (Euclid, Apollonius and others), who in fact considered the curves as sections of a cone by planes. This is why the second-order curves are commonly referred to as *conic sections.*

We take a right circular cone with vertex at S and the circle $ABCD$ as a transverse section (fig. 35). We suppose the cone to be generated indefinitely in both directions from the vertex, so that its surface is made up of two parts, $SABCD$ and $SA'B'C'D'$. We

take an arbitrary point M of the surface, and draw through M and the axis of the cone the axial section $MSCA$. We take a plane F through M perpendicular to the axial section, and rotate it about M so that it remains perpendicular to the axial section. In the posi-

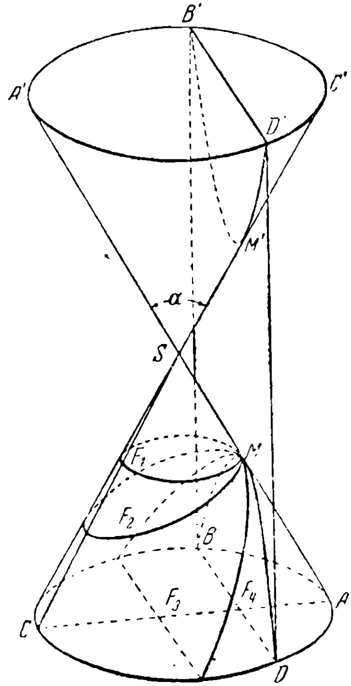


FIG. 35

tion F_1 of the plane F, in which it is perpendicular to the axis of the cone, its intersection with the conical surface is well known to consist of a circle. In any other position of the plane F for which its lesser angle with the axis of the cone is greater than half the angle of development α of the cone, the plane will only intersect the part $SABCD$ of the conical surface. The section here consists of an oval curve which is in fact an ellipse with a vertex at M. Figure 35 illustrates the ellipse corresponding to the position F_2 of the plane F. As the plane F approaches the position F_3 in which it is parallel to the generator SC (and forms an angle $\frac{1}{2}\alpha$ with the axis of the

cone) the ellipse lengthens and widens out, until at the position F_3 the section with the conical surface yields a parabola with vertex at M. On further rotation the plane F forms an angle less than $\frac{1}{2}\alpha$ with the axis of the cone and thus cuts both parts $SABCD$ and $SA'B'C'D'$ of the conical surface. One such position F_4 is shown in fig. 35, when the plane F forms an angle of 0° with the axis of the cone, i.e. when it is parallel to the axis. In this case the section of the conical surface by the plane yields a hyperbola, one branch MBD of which lies on $SABCD$, and the other $M'B'D'$ lies on $SA'B'C'D'$. The vertices of the hyperbola are at M and M'. The proofs of all these statements will be found in a comprehensive course of analytic geometry.

EXERCISES

Loci

1. Find the equation of the locus of a point which moves so that its distance from the Ox axis is five times its distance from Oy.

Ans. $y = \pm 5x$.

2. Find the equation of the locus of a point which moves so that its distance from Ox is equal to twice its distance from Oy plus 3 units of length.

Ans. $y = 2x + 3$, $x \geqslant 0$.

3. Find the equation of the locus of points equidistant from the points = $(2, -3)$ and $(3, 2)$.

Ans. $x + 5y = 0$.

4. A point moves on a plane so that its distance from the Oy axis always remains equal to its distance from the point $(5, 0)$. Find the equation of the curve described by the point.

Ans. $y^2 - 10x + 25 = 0$.

5. A point moves on a plane so that the square of its distance from the point $(0, 3)$ always remains equal to the cube of its distance

from the Oy axis. Find the equation of the curve described by the point.

Ans. $x^2+(y-3)^2=\pm x^3$.

6. Find the equation of the locus of a point which moves so that its distance from the straight line $x=3$ is equal to its distance from the point $(4, -2)$.

Ans. $y^2+4y-2x+11=0$.

7. Find the equation of the locus of a point which moves so that the slope of the straight line joining it to the origin is twice the slope of the straight line joining it to the point (a, a).

Ans. $xy-2ax+ay=0$.

8. A point moves on a plane so that its distance from the origin always remains equal to the slope of the straight line joining it to the origin. Find the equation of the curve described by the point.

Ans. $x^4+x^2y^2=y^2$.

The Circle

9. Find the equation of the circle of radius 4 with centre at the point $(3, -5)$.

Ans. $x^2+y^2-6x+10y+18=0$.

10. Find the equation of the circle of radius 2 and centre at the point $\left(-\frac{4}{5}, \frac{3}{5}\right)$.

Ans. $5x^2+5y^2+8x-6y-15=0$.

11. Find the points of intersection of Ox with the circle whose diameter is the straight line joining the points $(1, 2)$ and $(-3, -4)$.

Ans. $(-1\pm 2\sqrt{3}; 0)$.

12. Find the equation of the circle having as a diameter the intercept of the straight line $3x-4y+12=0$ contained between the co-ordinate axes.

Ans. $x^2+y^2+4x-3y=0$.

13. Find the equation of the circle having as a diameter the common chord of the circles $x^2+y^2+4x-4y-2=0$, x^2+y^2-2x+ $+2y-14=0$.

Ans. $x^2+y^2+2x-2y-6=0$.

14. Find the equation of the circle of radius a that touches Oy at the origin.

Ans. $x^2+y^2\pm 2ax=0$.

15. Find the centre and radius of the circle $x^2+y^2+2x+16y-$ $-42=0$.

Ans. $(-1\,,\,-8)$; $\sqrt{107}$.

16. Find the centre and radius of the circle $2x^2+2y^2+6x-$ $-3y-10=0$.

Ans. $\left(-\frac{3}{2}\,,\,\frac{3}{4}\right),\ \frac{5}{4}\sqrt{5}$.

17. Find the equation of the straight line through the centre of the circle $x^2+y^2-4x+2y-5=0$ and perpendicular to the straight line $x-2y+1=0$.

Ans. $2x+y-3=0$.

18. Find the equation of the circle passing through the points $(0\,,\,2)$, $(2\,,\,0)$ and $(0\,,\,0)$.

Ans. $x^2+y^2-2x-2y=0$.

19. Find the equation of the circle circumscribed about the triangle whose vertices are $(0\,,\,1)$, $(-2\,,\,0)$ and $(0\,,\,-1)$.

Ans. $2x^2+2y^2+3x-2=0$.

20. Find the equation of the circle circumscribing the isosceles triangle whose height is equal to 5 units of length and the base of which is the straight line joining the points $(-4\,,\,0)$, $(4\,,\,0)$.

Ans. $5x^2+5y^2\pm 9y-80=0$.

21. Find the equation of the circle circumscribing the triangle whose sides are the straight lines $x+2y-3=0$, $3x-y-2=0$, $2x-3y-6=0$.

Ans. $7x^2+7y^2-19x+11y-6=0$.

22. Find the equation of the circle touching the co-ordinate axes and passing through the point $(4, -2)$.

Ans. $x^2+y^2-4x+4y+4=0$.
$x^2+y^2-20x+20y+100=0$.

23. The centre of a circle that touches the co-ordinate axes lies on the straight line $3x-5y+15=0$. Find the equation of the circle.

Ans. $4x^2+4y^2-60x-60y+225=0$;
$64x^2+64y^2+240x-240y+225=0$.

24. A circle with a radius equal to 5 units of length passes through the points $(4, -2)$, $(5, -3)$. Find the equation of the circle.

Ans. $x^2+y^2-2x+12y+12=0$;
$x^2+y^2-16x-2y+40=0$.

25. The centre of a circle that passes through the points $(-2, 4)$ and $(-1, 3)$ lies on the straight line $2x-3y+2=0$. Find the equat tion of the circle.

Ans. $x^2+y^2+26x+16y-32=0$.

26. A circle touches Ox and passes through the points $(-1, 2)$ and $(6, 9)$. Find its equation.

Ans. $x^2+y^2-6x-10y+9=0$;
$x^2+y^2+18x-34y+81=0$.

27. Find the equation of the circle passing through the poin- $(-3, 4)$ and concentric with the circle $x^2+y^2+3x-4y-1=0$.

Ans. $x^2+y^2+3x-4y=0$.

28. A circle touches the straight lines $x-2=0$ and $x=6$, whilst its centre lies on the straight line $3x-y-6=0$. Find its equation.

Ans. $x^2+y^2-8x-12y+48=0$.

29. Show that, if the equations of two circles differ only in their free terms (i.e. the terms not containing the current co-ordinates), the circles are concentric.

30. Show that the locus of a point which moves so that the ratio of its distances from two fixed points is a constant k, where k is not equal to 1, is a circle.

31. Show that the locus of a point which moves so that the sum of the squares of its distances from the sides of a square is constant is a circle.

32. Show that the locus of a point which moves so that the square of its distance from a given point is proportional to its distance from a given straight line is a circle.

33. Two straight lines are drawn from the points $(-a, 0)$ and $(a, 0)$ such that the difference between their angles of inclination to Ox is equal to arc tan $\frac{1}{a}$. Show that the locus of their point of intersection is a circle.

The Ellipse

(It is assumed in all the examples of this section that the axes of symmetry of the ellipse coincide with the co-ordinate axes).

34. Find the lengths of the axes, the eccentricity and the co-ordinates of the foci of the ellipse $9x^2+25y^2=225$.

Ans. 10; 6; $\frac{4}{5}$; $(\pm 4, 0)$.

35. Find the lengths of the axes, the eccentricity and the co-ordinates of the foci of the ellipse $3x^2+4y^2=2$.

Ans. $\frac{2}{3}\sqrt{6}$; $\sqrt{2}$; $\frac{1}{2}$; $\left(\pm\frac{1}{6}\sqrt{6}, 0\right)$.

36. Find the co-ordinates of the vertices and foci, and the eccentricity, of the ellipse $4x^2+2y^2=1$.

Ans. $\left(\pm\frac{1}{2}, 0\right)$; $\left(0, \pm\frac{1}{2}\sqrt{2}\right)$; $\left(0, \pm\frac{1}{2}\right)$; $\frac{1}{2}\sqrt{2}$.

37. Find for what values of a and b the ellipse $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$ passes through the points $(2, 3)$, $(-1, -4)$.

Ans. $\frac{1}{7}\sqrt{385}$, $\frac{1}{3}\sqrt{165}$.

38. Find the equation of the ellipse with two vertices at $(\pm 4, 0)$ and one focus at $(2, 0)$.

Ans. $3x^2+4y^2=48$.

39. The co-ordinates of two vertices of an ellipse are $(\pm 6, 0)$, whilst the foci are at $(\pm 4; 0)$. Find the equation of the ellipse.

Ans. $5x^2+9y^2=180$.

40. Find the equation of the ellipse whose foci are at $(\pm 4, 0)$ and whose major axis is of length 10.

Ans. $9x^2+25y^2=225$.

41. Find the equation of the ellipse with foci at $(0, \pm 3)$ and major axis of length 12.

Ans. $4x^2+3y^2=108$.

42. Find the equation of the ellipse of which the minor axis is of length 6 and one focus has the co-ordinates $(-4, 0)$.

Ans. $9x^2+25y^2=225$.

43. Find the equation of the ellipse that passes through the points $(6, 4)$ and $(-8, 3)$.

Ans. $\frac{x^2}{100}+\frac{y^2}{25}=1$.

44. Find the equation of the ellipse that passes through the points $\left(\sqrt{11}, -\frac{10}{3}\right)$ and $\left(-3\sqrt{3}, 2\right)$.

Ans. $\frac{x^2}{36}+\frac{y^2}{16}=1$.

45. The major axis of an ellipse lies along Ox and its length is 6; the eccentricity is $\frac{1}{2}$. Find the equation of the ellipse.

Ans. $3x^2+4y^2=27$.

46. The eccentricity of an ellipse is $\frac{1}{3}$, whilst the ordinate of any point of the ellipse, whose abscissa is equal to the abscissa of a focus, is equal to 4 units of length. Find the equation of the ellipse.

Ans. $8x^2+9y^2=162$.

47. Find the equation of the ellipse of which the foci are $(0, \pm 5)$ and the eccentricity is $\frac{2}{3}$.

Ans. $36x^2+20y^2=1125$.

48. Find the equation of the ellipse of which the eccentricity is $\frac{1}{3}$ and the abscissa of one focus is $\frac{3}{2}$.

Ans. $8x^2+9y^2=162$.

49. Find the eccentricity of an ellipse, given that the ordinate of any point of it, whose abscissa is equal to the abscissa of a focus, has a length equal to $\frac{2}{3}$ of the minor semi-axis.

Ans. $\frac{\sqrt{5}}{3}$.

50. Find the eccentricity, given that the straight line joining the right-hand and upper vertices of an ellipse is parallel to the straight line joining the origin to the point of the ellipse whose ordinate is positive and whose abscissa is equal to the abscissa of the left-hand focus.

Ans. $\frac{\sqrt{2}}{2}$.

51. Find the points of intersection of the ellipse $\frac{x^2}{36}+\frac{y^2}{12}=1$ with the straight lines: (*a*) $y=4x$; (*b*) $x=6$; (*c*) $2x-y-9=0$; (*d*) $x-y+$ $+8=0$.

Ans. (*a*) $\left(\pm\frac{6}{7}, \pm\frac{24}{7}\right)$; (*b*) $(6, 0)$; (*c*) $(3, -3), \left(\frac{69}{13}, \frac{21}{13}\right)$; (*d*) they do not intersect.

52. A straight line moves so that its intercept contained between the co-ordinate axes preserves a constant length. Show that any point of the straight line describes an ellipse.

The Hyperbola

(It is assumed in all the problems in this section that the axes of symmetry of the hyperbola coincide with the co-ordinate axes).

53. Find the eccentricity, the co-ordinates of the foci and the equations of the asymptotes of the hyperbola

$$\frac{x^2}{4}-\frac{y^2}{25}=1.$$

Ans. $\frac{1}{2}\sqrt{29}$; $(\pm\sqrt{29}, 0)$; $5x\pm 2y=0$.

54. Find the eccentricity, the co-ordinates of the foci and the equations of the asymptotes of the hyperbola $4x^2-9y^2=36$.

Ans. $\frac{1}{3}\sqrt{13}$; $(\pm\sqrt{13}; 0)$; $2x\pm 3y=0$.

55. Find the eccentricity, the co-ordinates of the foci and the equations of the asymptotes of the hyperbola $\frac{y^2}{16}-\frac{x^2}{9}=1$.

Ans. $\frac{5}{4}$; $(0, \pm 5)$; $4x\pm 3y=0$.

56. Find the equation of the hyperbola with foci at $(\pm 4, 0)$ and transverse axis equal to 6.

Ans. $7x^2-9y^2=63$.

57. Find the equation of the hyperbola with asymptotes $y = \pm \frac{3}{5}x$ and foci at $(\pm 2, 0)$.

Ans. $153x^2 - 425y^2 = 450$.

58. Find the equation of the hyperbola with asymptotes $y = \pm \frac{3}{5}x$ that passes through the point $(2, 1)$.

Ans. $9x^2 - 16y^2 = 20$.

59. Find the equation of the rectangular hyperbola passing through the point $(3, -1)$.

Ans. $x^2 - y^2 = 8$.

60. Find the equation of the hyperbola, given that it passes through the points $\left(5, \frac{4}{3}\right)$ and $\left(\sqrt{34}, -\frac{5}{3}\right)$.

Ans. $\frac{x^2}{9} - y^2 = 1$.

61. Find the points of intersection of the hyperbola $\frac{x^2}{9} - \frac{y^2}{16} = 1$ with the straight lines:

(*a*) $2x + 3y = 0$; (*b*) $4x - 3y = 0$.

(*c*) $20x + 21y + 12 = 0$; (*d*) $2x - y - 3 = 0$.

Ans. (*a*) $\left(\pm 2\sqrt{3}, \pm \frac{4\sqrt{3}}{3}\right)$; (*b*) no intersection (the straight line is an asymptote); (*c*) $\left(\left(5, -\frac{16}{3}\right), \left(-\frac{15}{4}, 3\right)\right.$; (*d*) no intersection.

62. Find the points of intersection of the rectangular hyperbola $x^2 - y^2 = 16$ with the circle $x^2 + y^2 = 34$.

Ans. $(5, 3)$; $(-5, 3)$; $(-5, -3)$; $(5, -3)$.

63. Find the equation of the hyperbola with foci at $(0, \pm 3)$ and transverse axis equal to 4.

Ans. $5y^2 - 4x^2 = 20$.

64. Find the angle φ between the asymptotes and the transverse axis of a hyperbola, given that the distance of a vertex from the centre is $\frac{2}{3}$ the distance of a focus from the centre.

Ans. $\tan\varphi = \frac{1}{2}\sqrt{5}$.

65. Two vertices of an ellipse are located at the foci of a hyperbola, the vertices of which are at the foci of the ellipse. The equation of the ellipse is

$$\frac{x^2}{16}+\frac{y^2}{9}=1 .$$

Find the equation of the hyperbola.

Ans. $9x^2-7y^2=63$.

66. Show that the eccentricity of a rectangular hyperbola is equal to the ratio of the diagonal to the side of the square whose side is equal to an axis of the hyperbola.

67. Find the eccentricity, given that the angle between the asymptotes of a hyperbola is (*a*) 90°, (*b*) 60°.

Ans. (*a*) $\sqrt{2}$; (*b*) $\frac{2}{\sqrt{3}}$.

68. Show that the product of the distances of any point of a hyperbola from the asymptotes is constant.

69. Let a straight line be drawn through any point P of a hyperbola parallel to the transverse axis, and let it meet the asymptotes in Q and R. Show that $PQ \, . \, PR = a^2$.

70. A point moves on a plane so that the product of the slopes of the straight lines joining the point to $(-a, 0)$ and $(a, 0)$ is constant. Show that the point describes either an ellipse or a hyperbola.

The Parabola

71. Find the equation of the parabola, given that

(*a*) its axis of symmetry is Ox, its vertex is at the origin and the distance from the vertex to the focus is 6 units of length;

(*b*) it is symmetrical with respect to Ox, it passes through the point $(2, -5)$ and its vertex is at the origin;

(*c*) it is symmetrical with respect to Ox, it passes through $(-2, 4)$ and its vertex is at the origin;

(*d*) it is symmetrical with respect to Oy, its focus is $(0, 4)$ and its vertex is at the origin;

(*e*) it is symmetrical with respect to Oy, it passes through $(6, 3)$ and its vertex is at the origin;

(*f*) it is symmetrical with respect to Oy, it passes through $(-6, -3)$ and its vertex is at the origin.

Ans. (*a*) $y^2 = \pm 24x$; (*b*) $y^2 = \frac{25}{2}x$; (*c*) $y^2 = -8x$; (*d*) $x^2 = 16y$; (*e*) $x^2 = 12y$; (*f*) $x^2 = -12y$.

72. The vertex of a parabola is at (a, b) and its axis of symmetry is parallel to Ox. Find the equation of the parabola, given that its parameter is p.

Ans. $(y-b)^2 = \pm 2p(x-a)$.

Hint. See § 25.

73. A parabola passes through $(-1, -1)$ and has its vertex at $\left(-\frac{3}{2}, 2\right)$. Find its equation, given that its axis is parallel to Oy.

Ans. $12x^2 + 36x + y + 25 = 0$.

74. The vertex of a parabola is at $(2, 3)$, it passes through the origin, and its axis is parallel to Ox. Find the equation of the parabola.

Ans. $2y^2 - 12y + 9x = 0$.

75. Show that the equation $x = My^2 + Ny + P$ represents a parabola.

Hint. See § 25.

76. Find the co-ordinates of the vertex and focus, and the equations of the axis and directrix of the parabola $y^2+4y-6x+$ $+7=0$.

Ans. $\left(\frac{1}{2}, -2\right)$; $(2, -2)$; $y+2=0$; $x+1=0$.

77. Find the co-ordinates of the vertex and focus, and the equations of the axis and directrix of the parabola $4x^2+4x+3y-2=0$.

Ans. $\left(-\frac{1}{2}, 1\right)$; $\left(-\frac{1}{2}, \frac{13}{16}\right)$; $2x+1=0$; $16y-19=0$.

78. Find the equation of the parabola of parameter p whose focus is at the origin and whose axis lies along Ox.

Ans. $y^2=\pm 2px+p^2$.

79. Find the equation of the parabola of parameter p whose axis and directrix lie along Ox and Oy respectively.

Ans. $y^2=\pm 2px-p^2$.

80. Find the equation of the parabola whose vertex is at $(3, 2)$ and focus at $(5, 2)$.

Ans. $y^2-4y-8x+28=0$.

81. Find the equation of the parabola whose vertex is at $(-1, -2)$ and focus at $(-1, -4)$.

Ans. $x^2+2x+8y+17=0$.

82. Find the equation of the parabola whose focus is at $(2, -1)$ and whose directrix lies along $y-4=0$.

Ans. $x^2-4x+10y-11=0$.

83. Find the equation of the parabola whose vertex is at $(-2, -5)$ and whose directrix is $x-3=0$.

Ans. $y^2+10y+20x+65=0$.

84. The directrix of a parabola is $y+4=0$ and the vertex is at $(5, -2)$. Find the equation of the parabola.

Ans. $x^2-10x-8y+9=0$.

85. Find the co-ordinates of the vertex and focus, and the equations of the axis and directrix, for the following parabolas:

(*a*) $x^2-8x-16y+32=0$;
(*b*) $x^2-8x+8y+8=0$;
(*c*) $3x^2-2x+y+5=0$;
(*d*) $4y^2-8y-13x-12=0$;
(*e*) $y=10x-x^2$;
(*f*) $y^2-4x+8=0$.

Ans. (*a*) $(4, 1)$; $(4, 5)$; $x=4$; $y+3=0$;

(*b*) $(4, 1)$; $(4, -1)$; $x=4$; $y-3=0$;

(*c*) $\left(\frac{1}{3}, -4\frac{2}{3}\right)$; $\left(\frac{1}{3}, -4\frac{3}{4}\right)$; $3x-1=0$; $12y+55=0$;

(*d*) $\left(-\frac{16}{13}, 1\right)$; $\left(-\frac{87}{208}, 1\right)$; $y=1$; $208x+425=0$;

(*e*) $(5, 25)$; $\left(5, 24\frac{3}{4}\right)$; $x=5$; $4y-101=0$;

(*f*) $(2, 0)$; $(3, 0)$; $y=0$; $x-1=0$.

86. A bridge has a parabolic arch. Find the parameter p of the parabola, given that the span of the arch is 24 m and its height 6 m.

Ans. $p=12$.

87. Figure 36 shows a longitudinal section through a parabolic mirror. Using the data given, find the abscissa of the focus.

Ans. $x=5{\cdot}625$.

88. Assuming that the cable joining the points M and N (fig. 37) has a parabolic form, find the equation of the parabola with $p=0{\cdot}1$, $h=1$ and $l=10$.

Ans. $y=-0{\cdot}02\,(1+\sqrt{11})x+0{\cdot}002\,(6+\sqrt{11})\;x^2$.

Hint. Take the equation of the parabola in the form $y=Ax^2+$ $+Bx+C$. The ordinate of the vertex is given by

$$b=\frac{4AC-B^2}{4A}.$$

Also, the parabola passes through the points M and N, whose co-ordinates are known. Three equations in the coefficients A, B, C can be formed from these conditions.

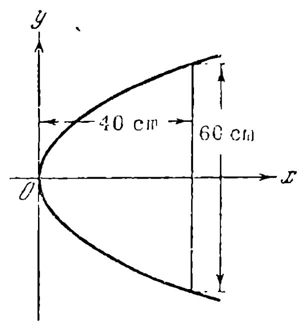


FIG. 36

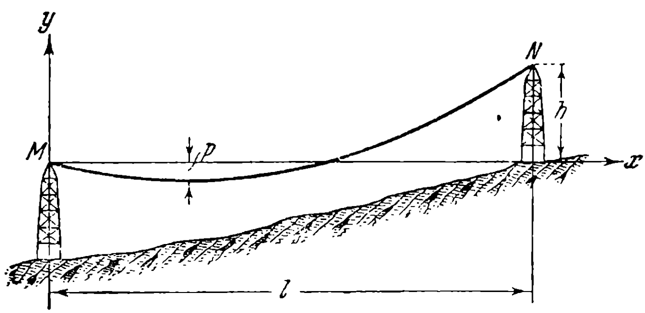


FIG. 37

89. Find the points of intersection of two parabolas having a common vertex at the origin and foci at (2, 0) and (0, 2).

Ans. (0, 0); (8, 8).

90. Find the points of intersection of the ellipse $\frac{x^2}{100}+\frac{y^2}{64}=1$ with the parabola whose vertex is at the centre of the ellipse and whose focus coincides with the right-hand focus of the ellipse.

Ans. $\left(\frac{5}{2},\quad \pm 2\sqrt{15}\right)$.

91. An equilateral triangle is inscribed in the parabola $y^2=2px$. One vertex of the triangle coincides with the vertex of the parabola. Find the length of side of the triangle.

Ans. $4p\sqrt{3}$

92. Show that the locus of a point P, which moves so that its ordinate is equal to the abscissa of the point of intersection of the straight line from P to the origin with the straight line $y=a$, is a parabola.

93. Find the locus of a point which moves so that its shortest distance from a given circle is equal to its distance from a fixed diameter of the circle.

Ans. Two parabolas.

94. AB and CD are perpendicular diameters of a circle. From a variable point M of the circle we draw AM and BM. Let AM meet CD at N; we draw through N a straight line parallel to AB to meet BM in P. Show that the locus of P is a parabola.

95. The vertex of a triangle having a fixed base varies so that the sum of the tangents of the base angles of the triangle remains constant. Show that the vertex describes a parabola.

96. Find the locus of the centres of circles which pass through a fixed point and touch a fixed straight line.

Ans. A parabola.

97. Find the locus of the centres of circles which touch a fixed circle and a fixed straight line.

Ans. A parabola.

PART II

ELEMENTS OF DIFFERENTIAL CALCULUS

CHAPTER 4

THE THEORY OF LIMITS

§ 27. Some relationships between absolute numerical values. We constantly find ourselves obliged in mathematical analysis to consider relationships between the absolute values of various expressions. We shall recall the relevant basic formulae in the present article.

The absolute value of a number a is defined as a itself if the number is positive or zero, or as $-a$ if a is negative. The absolute value of the number a is written $|a|$.

Thus $|a|=a$ if $a\geqq 0$; $|a|=-a$, if $a<0$. $|5|=5$; $|-5|=-(-5)=5$.

1. *The absolute value of an algebraic sum is less than or equal to the sum of the absolute values of the terms:*

$$|a+b+c+\ldots+v|\leqslant |a|+|b|+|c|+\ldots+|v|.$$

We illustrate the truth of this proposition with the following examples:

$$|3+5+8|=|16|=16\ ;$$
(1)
$$|3|+|5|+|8|=3+5+8=16\ ;$$

whence

(2) $$|3+5+8|=|3|+|5|+|8|.$$

$$|-3-5-8|=|-16|=16\ ;$$

$$|-3|+|-5|+|-8|=3+5+8=16\ ;$$

whence

(3) $$|-3-5-8|=|-3|+|-5|+|-8|.$$

$$|3-5+8|=|6|\ ;$$

$$|3|+|-5|+|8|=3+5+8=16\ ;$$

whence

$$|3-5+8| < |3| + |-5| + |8|.$$

Evidently, we have the sign of equality in (1) when all the terms have the same sign, and the sign of inequality when the terms have different signs.

Thus in the general case, i.e. when the signs of the terms of the algebraic sum may be the same or different, we can only assert that *the absolute value of the algebraic sum does not exceed* (i.e. is either less than, or equal to) *the arithmetic sum of the absolute values of the terms.*

2. *The absolute value of the product of any number of factors is independent of the signs of the individual factors. Hence we always have*

$$|a \cdot b \cdot c \ldots v| = |a| \cdot |b| \cdot |c| \ldots |v|.$$

3. *The absolute value of a quotient is independent of the signs of the numerator and denominator. Hence we always have*

$$\left|\frac{a}{b}\right| = \frac{|a|}{|b|}.$$

§ 28. Variables and constants. Simple observation of the world about us forces us to distinguish between two types of quantities: constants and variables. For instance, the height above ground of a stone thrown upwards is a variable, whilst the volume of the stone is a constant.

Strictly speaking, the volume of our stone changes for various reasons, e.g. because of variation of the atmospheric temperature. The change is so insignificant, however, when a stone is thrown upwards, that we assume the volume constant in practice. It is doubtful if anyone buying material stops to consider the change in length of the metre stick used for measuring the material, though in reality the length of the stick is constantly fluctuating due to a number of factors, e.g. the humidity and temperature of the atmosphere, which affect the substance of which the stick is composed.

In fact, practical experience compels us to distinguish between constants and variables.

We are not concerned in mathematics with the physical nature of a magnitude but only with the number that expresses it.

A mathematical quantity is called a variable if it can take different numerical values in the conditions of the problem concerned.

The hypotheses of each particular problem determine which of the quantities concerned are variables and which are constants.

For instance, if the vertex of a triangle moves along a straight line parallel to the base, the angles and sides of the triangle are variables, whilst the base, height and sum of the angles are constants.

If the vertex were to move along a straight line not parallel to the base, the height and area of the triangle would be variables and only the base and sum of the angles would be constant.

There are, however, quantities that are constant whatever the problem; constants of this sort include, for instance, the sum of the angles of a triangle, the ratio of the circumference to the diameter of a circle (the number π), the numbers 5, -1, $\sqrt{3}$, etc.

Constants are usually denoted by the first letters of the alphabet, a, b, c, ..., and variables by the last: x, y, z, ...

We shall always assume below that the values considered for the quantities are real numbers.

§ **29. Infinitesimals.** 1. Quantities that vary in such a way that their numerical values indefinitely approach zero play an essential part in mathematical analysis. Variables of this type are encountered so often that they are distinguished by the special name of *infinitesimals*, whilst their properties are the subject of detailed investigation.

Definition. A variable α is called an infinitesimal if, in the course of its variation, its absolute value becomes and then remains less than any previously assigned positive number ε, however small:

$$|\alpha| < \varepsilon\,.$$

We shall consider some examples of infinitesimals.

Example 1. We take a pendulum which starts to oscillate after being displaced from the equilibrium position (fig. 38). We define

the position of the pendulum by the angle α that it forms with the vertical (equilibrium position). We consider α positive or negative according to whether the pendulum is to the right or left of the vertical.

The amplitude of the oscillation will steadily decrease due to the resistance of the medium. Thus, whatever the positive number

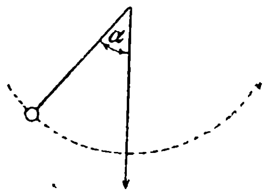


FIG. 38

ε that we assign, the absolute value of the deviation α will become and then remain less than ε.

In other words, α is an infinitesimal and in the course of its variation it takes negative and zero, as well as positive, values.

Example 2. We show that the variable $y=x^3$ is an infinitesimal when x approaches zero indefinitely.

We assign some positive number ε, say $\varepsilon=0{\cdot}001$. The inequality

$$|y|<0\cdot 001$$

or what amounts to the same thing,

$$|x^3|<0\cdot 001$$

will be satisfied provided that x, as it approaches zero, becomes less in absolute value than $\sqrt[3]{0{\cdot}001}=0{\cdot}1$, i.e.

$$|x|<0\cdot 1\,.$$

The inequality $|y|<0{\cdot}001$ will clearly remain valid as x further approaches zero.

Now suppose we take a smaller positive number ε, say $\varepsilon=0{\cdot}000001$. The inequality

$$|y|<0{\cdot}000001$$

or what amounts to the same thing,

$$|x^3| < 0{\cdot}000001$$

is valid provided that the absolute value of x becomes less than $\sqrt[3]{0{\cdot}000001} = 0{\cdot}01$, i.e.

$$|x| < 0{\cdot}01 .$$

The inequality $|y| < 0{\cdot}000001$ will clearly remain valid as x further approaches zero.

The situation will be similar whatever the previously assigned ε: provided only that the absolute value of x becomes less than $\sqrt[3]{\varepsilon}$, i.e.

$$|x| < \sqrt[3]{\varepsilon},$$

we shall have

$$|x|^3 < \varepsilon,$$

and this latter inequality will remain valid as x further approaches zero.

The variable y thus satisfies the test defining an infinitesimal as $x \to 0$.

Example 3. We show that the ratio $\frac{1}{x}$ is an infinitesimal on indefinite increase of x, or as we usually say, as x tends to $+\infty$ ($x \to +\infty$).

We notice first of all that, since x increases indefinitely, we need only consider positive values of x, so that

$$\left|\frac{1}{x}\right| = \frac{1}{x} .$$

We take $\varepsilon = \frac{1}{1\,000\,000}$. The inequality

$$\frac{1}{x} < \frac{1}{1\,000\,000}$$

is valid as soon as our increasing x becomes greater than 1 000 000, and evidently remains valid on further increase of x.

In general, whatever the positive ε, the inequality

$$\frac{1}{x} < \varepsilon$$

becomes valid as soon as x becomes greater than $\frac{1}{\varepsilon}$, and remains valid on further increase of x.

Example 4. The ratio $-\frac{1}{x}$ is also an infinitesimal on indefinite increase of x.

In fact, the definition of an infinitesimal α says that the absolute value of α, and not α itself, must become and then remain less than any previously assigned positive number ε; so that it is of no consequence whether α takes only positive or only negative values, or both positive and negative values, during its variation.

Whilst $x \to +\infty$, we have $\left|-\frac{1}{x}\right| = \frac{1}{x}$. And since we have shown that $\frac{1}{x}$ is an infinitesimal for $x \to +\infty$, it follows that $-\frac{1}{x}$ is likewise an infinitesimal.

Example 5. We again consider $\frac{1}{x}$ and show that it is an infinitesimal when the absolute value of x increases indefinitely and x remains negative. This is the case we have in mind when we say that "x tends to minus infinity" ($x \to -\infty$).

Let $x = -z$. When $x \to -\infty$, $z \to +\infty$. Moreover, $\left|\frac{1}{x}\right| = \frac{1}{|x|} = \frac{1}{-x} = \frac{1}{z}$. We proved in example 3 that $\frac{1}{z}$ is an infinitesimal for $z \to +\infty$. Hence $\frac{1}{x}$ is also an infinitesimal, when $x \to -\infty$.

It must be remarked that the term "infinitesimal" is extremely unfortunate and often leads to misunderstandings: it suggests an

idea of the size of a quantity, whereas it must in fact be looked on as merely descriptive of the nature of the variation of a quantity.

The term is preserved for historical reasons. It dates from a time when quite a different meaning from the present one was attached to the concept of infinitesimal.

2. If cold water is heated to boiling-point, the water temperature is a variable that gradually increases up to the instant when the water starts to boil. Throughout the time that the water boils, its temperature remains constant at 100°C. If the water is allowed to cool, its temperature starts to fall, i.e. becomes a variable again.

Here we have an example of a quantity that first varies, then becomes constant, then again varies. It seems natural to consider such a quantity as in general a variable.

If we pay attention only to the numerical value of the temperature and disregard the physical significance of the process concerned, we are presented with a mathematical quantity that it also seems natural to regard as a variable.

Furthermore, it is often convenient to regard a mathematical quantity that always remains constant, as a variable that only takes a single value.

If this single numerical value is zero, the constant in question belongs to the class of infinitesimals, in as much as it satisfies the definition of infinitesimal: whatever positive ε we may be given, we always have

$$|0| = 0 < \varepsilon .$$

A constant a differing from zero is not an infinitesimal, since no constant $a = 0$ exists whose absolute value is less than any positive ε, however small. To see this, we only need to take as ε some positive number less than $|a|$.

§ 30. Basic properties of infinitesimals. 1. First property of infinitesimals. *The algebraic sum of any (constant) number of infinitesimals is itself an infinitesimal.*

Proof. For simplicity, we prove the theorem for the sum of two infinitesimals, $\alpha + \beta = \sigma$.

Let ε be any positive number, however small. We want to show that a point is reached in the course of the variation of σ at and after which we have

$$|\sigma| < \varepsilon .$$

Since both α and β are infinitesimals, which, in general, vary in different ways, their absolute values must sooner or later become, and then remain, less than $\frac{1}{2}\varepsilon$. Hence a point will be reached at which we can write for both variables,

$$|\alpha| < \tfrac{1}{2}\varepsilon \quad \text{and} \quad |\beta| < \tfrac{1}{2}\varepsilon .$$

Now:

$$|\sigma| = |\alpha+\beta| \leqslant |\alpha| + |\beta| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon ,$$

i.e.

$$|\sigma| < \varepsilon ,$$

which is what we wanted to prove.

The proof remains the same for the sum of any number of infinitesimals, provided the number of terms in the sum is constant.

Note. The property proved does not extend to the case when the number of terms in the sum of infinitesimals increases indefinitely instead of remaining constant as each term approaches zero. It may happen in this case that the sum is not an infinitesimal.

Suppose, for instance, that a straight line of unit length is divided into n equal parts. Each part is equal to $\frac{1}{n}$, whilst the sum of all the parts is equal to unity. We now increase indefinitely the number n of parts. Each part, equal to $\frac{1}{n}$, is now an infinitesimal (see example 3, § 29). But whatever the value of n, the sum of all the parts is always equal to unity and is not an infinitesimal.

2. *A variable is said to be bounded if, during its variation, its absolute value never exceeds a certain positive number.* For instance, $\sin x$ is bounded when x varies, since we always have

$$|\sin x| \leqslant 1 .$$

Second property of infinitesimals. *The product of a bounded quantity and an infinitesimal is an infinitesimal.*

Proof. Let y be a bounded quantity and α an infinitesimal, and let z denote their product:

$$z = y\alpha .$$

We want to show that z is an infinitesimal, i.e. there is a point in the variation of z at and after which we have

$$|z| < \varepsilon ,$$

where ε is any previously assigned positive number, however small.

Since y is bounded, there exists a positive number N such that $|y| < N$. Since α is an infinitesimal, there is a point at and after which we can write $|\alpha| < \dfrac{\varepsilon}{N}$. We now have:

$$|z| = |y \cdot \alpha| = |y| \cdot |\alpha| < N \frac{\varepsilon}{N} = \varepsilon .$$

Constants and infinitesimals are clearly bounded, so that the property proved leads to the corollaries:

COROLLARY 1. *The product of a constant and an infinitesimal is an infinitesimal.*

COROLLARY 2. *The product of two infinitesimals is an infinitesimal.*

Finally, it is easy to prove the following corollary on the basis of the second property of infinitesimals:

COROLLARY 3. *The product of any constant number of factors, of which at least one is an infinitesimal whilst the remainder are bounded (or, in particular, rare constants), is an infinitesimal.*

§ **31. The limit of a variable.** Variables will have been encountered in elementary geometry and algebra which vary in such a way

as to approach a definite number, with the difference between the number and the variable becoming as small as desired.

For instance, the difference between the area of a circle and the area of the inscribed regular polygon becomes as small as desired on doubling indefinitely the number of sides of the polygon.

DEFINITION. *The number A is said to be the limit of a variable y if the difference*

$$y-A=\alpha$$

is an infinitesimal.

Or we can say that *the variable y tends to the limit A (A being a constant number) if the difference $y-A=\alpha$ is an infinitesimal.*

We indicate the fact that A is the limit of y by writing

$$\lim y=A\,,$$

or else we say that y tends to A and write $y\to A$.

We have from the equation $y-A=\alpha$:

$$y=A+\alpha\,,$$

i.e. *a variable y with the limit A can be expressed as the sum of two terms: the constant A* (i.e. the limit) *and an infinitesimal α.*

Conversely, *if a variable y is the sum of a number A and an infinitesimal α, A is the limit of y.*

For the equation $y=A+\alpha$ gives us $y-A=\alpha$, where α is an infinitesimal. It now follows from the definition that A is the limit of y.

It follows from the definition of limit that:

1. *The limit of an infinitesimal α is zero.* For the difference $\alpha-0=\alpha$ is an infinitesimal.

2. *If* $\lim \alpha=0$, *α is an infinitesimal.*

For it follows from the equation $\lim \alpha=0$ that the difference $\alpha-0$ is an infinitesimal. But $\alpha-0=\alpha$, so that α is an infinitesimal.

3. We have agreed to consider a constant c as a variable y taking the unique value c; recalling that we placed a constant equal to zero in the class of infinitesimals in 2. in § 29, we get

$$y-c=c-c=0$$

i.e. $y-c$ is an infinitesimal; whence it follows that $\lim y=c$, i.e. $\lim c=c$.

Thus the *limit of a constant is equal to the constant itself.*

Using the definition of infinitesimal, we can define the limit of a variable in the following alternative way:

The number A is said to be the limit of the variable y if the difference $y-A$ has an absolute value that becomes and afterwards remains, during the course of the variation of y, less than any previously assigned positive number ε however small:

$$|y-A|<\varepsilon .$$

On no account should it be thought that every variable has a limit.

Take, for instance, $y=\sin x$, varying in accordance with the variation of x. Let x increase through all real values ($x\to+\infty$). The variable $y=\sin x$ now oscillates an indefinite number of times between the values -1 and $+1$. Due to this, there is no number A for which the difference $y-A$ becomes an infinitesimal. Hence our variable $y=\sin x$ has no limit.

We shall encounter further examples of variables without limits a little later, in § 34.

4. In mathematical analysis we constantly encounter variables whose variation is due to that of other variables. Let y vary in accordance with the variation of x. If we speak about the limit of y as x approaches indefinitely (tends to) the number a ($x\to a$), *the number a is itself excluded from the values taken by x.* This condition is of great importance in the theory of limits.

5. We now explain the definition of the limit of a variable with the aid of some examples.

(I) We consider the variation of

$$y=\frac{x^2+3x-4}{x-1},$$

as x varies, on the assumption that $x\to1$.

On assigning to x the sequence of values

$$x=1\cdot1\,,\quad x=1\cdot01\,,\quad x=1\cdot001\,,\quad x=1\cdot0001\,,\ldots$$

we find for the corresponding values of y:

$$y=5\cdot 1\,,\quad y=5\cdot 01\,,\quad y=5\cdot 001\,,\quad y=5\cdot 0001\,,\ldots$$

A consideration of these values suggests that the limit of y as $x\to 1$ is 5; but to prove this, we have to show that

$$\frac{x^2+3x-4}{x-1}-5=\alpha$$

is an infinitesimal. We prove that this is in fact the case.

We do this by transforming the expression for α as follows: writing

$$\alpha=\frac{x^2+3x-4}{x-1}-5=\frac{x^2+3x-4-5x+5}{x-1}$$

$$=\frac{x^2-2x+1}{x-1}=\frac{(x-1)^2}{x-1}=x-1$$

(it is permissible to cancel by $x-1$ because the value $x=1$ is excluded from the values taken by x as $x\to 1$, so that $x-1$ cannot vanish as $x\to 1$).

Obviously, $x-1=\alpha$ is an infinitesimal for $x\to 1$. And since the difference

$$\frac{x^2+3x-4}{x-1}-5=\alpha$$

is an infinitesimal, it follows from the definition of limit that 5 is the limit of $y=\dfrac{x^2+3x-4}{x-1}$ 25 $x\to 1$.

The result obtained is written

$$\lim_{x\to 1}\frac{x^2+3x-4}{x-1}=5\,.$$

(II) We show that the limit of

$$y=\frac{x}{1-\sqrt{1-x}}$$

as $x\to 0$ is equal to 2.

To show that 2 is the limit of y, we have to show that the difference

$$y-2=\frac{x}{1-\sqrt{1-x}}-2=\alpha$$

is an infinitesimal. We use the following transformations:

$$\alpha=\frac{x}{1-\sqrt{1-x}}-2=\frac{x(1+\sqrt{1-x})}{(1-\sqrt{1-x})(1+\sqrt{1-x})}-2$$
$$=\frac{x(1+\sqrt{1-x})}{x}-2=\sqrt{1-x}-1\,;$$

(cancelling by x is permissible because x cannot vanish as $x\to 0$ by the condition of 4. above). Hence

$$\alpha=\sqrt{1-x}-1\,.$$

As x tends to zero, $\sqrt{1-x}$ tends to 1, whilst the difference $\sqrt{1-x}-1=\alpha$ tends to zero. This means that α is an infinitesimal. Thus

$$\lim_{x\to 0}\frac{x}{1-\sqrt{1-x}}=2\,.$$

These examples will have shown that difficulties can be encountered when finding limits; the difficulties may in fact be considerable, as we shall see more than once in what follows.

The following point may be noted here: if we find by working out a series of numerical values for a variable that it approaches a certain number in the course of its variation (cf. example 1), this is still not sufficient for asserting that the number is the limit of the variable: we have to prove that the difference between the number and the variable is an infinitesimal. Thus the recurring decimal 0·989898... always increases and approaches closer to unity as we increase the number of decimal places. But we are mistaken in concluding from this that 1 is the limit of the decimal because, however many places we take, the difference between 1 and the decimal is always greater than 1/99: as we know from arithmetic, the limit of the decimal is 98/99.

The following article presents a number of theorems that facilitate finding the limits of variables. It is by no means true, however, that the limit of a variable can always be found by direct application of these theorems. Generally speaking, the problem of limit-finding is one of considerable difficulty.

§ **32. Basic limit theorems.** Preliminary remark. The theorems proved below *assume the existence of limits* for all the variables concerned.

THEOREM 1. *The limit of the algebraic sum of a constant number of variables is equal to the algebraic sum of the limits of the terms.*

Proof. For brevity, we prove the theorem for two variables, $y+z$. The proof is similar, whatever the constant number of terms.

Let $\lim y=A$ and $\lim z=B$. Then we have by the definition of limit (§ 31):

$$y=A+\alpha\,, \qquad z=B+\beta\,,$$

where α and β are infinitesimals.

Hence

$$y+z=(A+B)+(\alpha+\beta)\,.$$

In the equation obtained, the sum $y+z$ is a variable, $A+B$ is a constant (number), and $\alpha+\beta$ is an infinitesimal, being the sum of infinitesimals (§ 30, property 1). Thus the variable $y+z$ consists of the sum of a number $A+B$ and an infinitesimal, whence it follows at once that $A+B$ is the limit of $y+z$ (cf. § 31):

$$\lim\,(y+z)=A+B$$

or

$$\lim\,(y+z)=\lim\,y+\lim\,z\,.$$

The theorem is proved.

THEOREM 2. *The limit of the product of a constant number of variables is equal to the product of their limits.*

Proof. For brevity, we prove the theorem for two variables y and z.

Let lim $y=A$ and lim $z=B$. From the definition of limit (§ 31):

$$y=A+\alpha\,,\qquad z=B+\beta\,,$$

where α and β are infinitesimals. Hence

$$yz=(A+\alpha)\,(B+\beta)=AB+A\beta+B\alpha+\alpha\beta\,.$$

Each of the terms $A\beta$, $B\alpha$, $\alpha\beta$ is an infinitesimal (§ 30, property 2), so that their sum also is an infinitesimal (§ 30, property 1). The product AB is a number.

Thus

$$\lim\,(yz)=AB$$

or

$$\lim\,(yz)=\lim y\cdot\lim z\,.$$

THEOREM 3. *The limit of the quotient of two variables is equal to the quotient of their limits, provided the limit of the denominator is not zero.*

Let y and z be the two variables, and let lim $y=A$, lim $z=B$, where $B\neq0$.

The theorem says that, if these conditions are satisfied, firstly the limit of z exists, and secondly, this limit is equal to the quotient of the limits of y and z. The proof of the first assertion is fairly complicated and we shall therefore omit it.

Assuming that the limit of the quotient exists, the second assertion can be proved as follows:

Let $\dfrac{y}{z}=v$, whence $y=z.v$. Since lim z and lim v exist, we have by theorem 2:

$$\lim\,(z\cdot v)=\lim z\cdot\lim v$$

so that
$$\lim y=\lim z\cdot\lim v,$$

i.e.
$$A=B\cdot\lim v\,.$$

Since $B\neq0$, we can divide both sides of this last equation by B to obtain

$$\lim v = \frac{A}{B}$$

or

$$\lim \frac{y}{z} = \frac{\lim y}{\lim z},$$

which is what we wished to prove.

We finally state a self-evident theorem about limits, the proof of which is omitted.

THEOREM 4. *If the variables u, y, v satisfy the inequality $u < y < v$ in the course of their variation and u and v have the common limit L, y tends to the same limit L.*

We take some examples of applications of these theorems.

Example 1. Find $\lim\limits_{x\to 1} \dfrac{x^2-2x+5}{x^2+7}$.

Solution. Since (theorems 1 and 2)

$$\lim_{x\to 1}(x^2+7) = \lim_{x\to 1} x^2 + \lim_{x\to 1} 7 = (\lim_{x\to 1} x)^2 + 7 = 1^2 + 7 = 8 \neq 0,$$

we find on applying theorems 1, 2 and 3:

$$\lim_{x\to 1}\frac{x^2-2x+5}{x^2+7} = \frac{\lim\limits_{x\to 1}(x^2-2x+5)}{\lim\limits_{x\to 1}(x^2+7)} = \frac{\lim\limits_{x\to 1} x^2 - \lim\limits_{x\to 1}(2x) + \lim\limits_{x\to 1} 5}{8}$$

$$= \frac{(\lim x)^2 - \lim 2 \cdot \lim x + 5}{8} = \frac{1^2 - 2\cdot 1 + 5}{8} = \frac{1}{2}.$$

Example 2. Find $\lim\limits_{x\to 0} \dfrac{2x^4+3x^3}{x^3}$.

Solution. Here the limit of the denominator is zero, and theorem 3 cannot be applied directly. In cases like this we carry out preliminary transformations of the expression in order to reduce it to a form to which the theorems regarding limits can be applied.

Since $x=0$ is excluded from the values taken by x as $x\to 0$, we can cancel the expression in our example by x^3, to give

$$\lim_{x\to 0}\frac{2x^4+3x^3}{x^3}=\lim_{x\to 0}(2x+3)=\lim_{x\to 0}(2x)+\lim_{x\to 0}3$$

$$=\lim_{x\to 0}2\cdot\lim_{x\to 0}x+3=2\cdot 0+3=3\,.$$

Example 3. Find $\lim\limits_{x\to 1}\dfrac{x^2+3x-4}{x-1}$.

In example 1 (§ 31), *starting from the assumption* that the variable $y=\dfrac{x^2+3x-4}{x-1}$ has the number 5 as its limit, *we proved* that 5 is in fact the limit of this variable. We now set the problem differently: we have *to find* the unknown limit of $\dfrac{x^2+3x-4}{x-1}$ as $x\to 1$.

Solution. Since we cannot apply theorem 3 directly for the solution of this problem (the limit of the denominator is zero), we carry out preliminary transformations. We have

$$\lim_{x\to 1}\frac{x^2+3x-4}{x-1}=\lim_{x\to 1}\frac{x^2+4x-x-4}{x-1}=\lim_{x\to 1}\frac{x(x+4)-(x+4)}{x-1}$$

$$=\lim_{x\to 1}\frac{(x+4)\,(x-1)}{x-1}=\lim_{x\to 1}(x+4)=5\,.$$

(Cancelling by $x-1$ is permissible because $x=1$ is excluded from the values taken by x.)

§ 33. Infinitely large quantities. 1. DEFINITION. *A variable y is said to be infinitely large if there is a point in its variation at and after which its absolute value is greater than any previously assigned positive number N* (however large), i.e.

$$|y|>N\,.$$

An infinitely large quantity is thus certainly not bounded.

The variable $y=\tan x$ where $x\to\frac{1}{2}\pi$ can serve as an example of an infinitely large quantity.

If x tends to $\frac{1}{2}\pi$, remaining less than $\frac{1}{2}\pi$ (and greater than zero), tan x increases indefinitely whilst remaining positive. If $x \to \frac{1}{2}\pi$, remaining $> \frac{1}{2}\pi$ (and $< \pi$), tan x remains negative whilst its absolute value increases indefinitely. Finally, if x tends to $\frac{1}{2}\pi$ and meantime takes values less than as well as greater than $\frac{1}{2}\pi$, tan x takes positive as well as negative values whilst its absolute value increases indefinitely.

The term "infinitely large quantity" is as unfortunate as the term "infinitesimal" and we can repeat here the remarks made in 1. in § 29 regarding the latter term.

2. An infinitely large quantity y cannot tend to a limit A, since as y varies its absolute value becomes greater than the absolute value $|A|$ of any A, whilst $y-A$, which increases in absolute value on further variation of y, cannot be an infinitesimal.

Nevertheless we often find it convenient to speak of the limit of an infinitely large positive quantity y as being plus infinity ($\lim y = +\infty$), the limit of an infinitely large negative quantity as being minus infinity ($\lim y = -\infty$), and the limit of an infinitely large quantity y which does not preserve a definite sign as y varies as simply ∞ ($\lim y = \infty$).

Thus when $x \to \frac{1}{2}\pi$, remaining $< \frac{1}{2}\pi$, $\lim\limits_{x \to \frac{1}{2}\pi} \tan x = +\infty$, whilst when $x \to \frac{1}{2}\pi$, remaining $> \frac{1}{2}\pi$, $\lim\limits_{x \to \frac{1}{2}\pi} \tan x = -\infty$; finally, when $x \to \frac{1}{2}\pi$ in an arbitrary manner, $\lim\limits_{x \to \frac{1}{2}\pi} \tan x = \infty$*.

* This is the sense in which we should understand the not-strictly-accurate expression: $\tan \frac{1}{2}\pi = \pm\infty$, which is found in some trigonometry textbooks.

§ 34. The connexion between infinitely large quantities and infinitesimals. 1. The relationship between infinitely large quantities and infinitesimals is shown by the following two theorems:

THEOREM 1. *If y is an infinitely large quantity, $\frac{1}{y}$ is an infinitesimal.*

Proof. Let ε be a positive number as small as we please. Since y is an infinitely large quantity there will be a point in its variation at and after which we have

$$|y| > \frac{1}{\varepsilon};$$

whence it follows from the theory of inequalities that we shall also have

$$\frac{1}{|y|} < \varepsilon, \quad \text{or} \quad \left|\frac{1}{y}\right| < \varepsilon,$$

which implies that $\frac{1}{y}$ is an infinitesimal.

THEOREM 2. *If y is an infinitesimal, $\frac{1}{y}$ is an infinitely large quantity* (we assume here that y does not take values equal to zero).

Proof. We take a positive number N as large as we please. Since y is an infinitesimal, as from a certain point in its variation y will satisfy

$$|y| < \frac{1}{N};$$

we shall now also have

$$\frac{1}{|y|} > N \quad \text{or} \quad \left|\frac{1}{y}\right| > N,$$

which implies that $\frac{1}{y}$ is an infinitely large magnitude.

Theorem 1 is often used for finding the limits of variables.

Example 1. Find $\lim\limits_{x \to +\infty} \frac{x+1}{x}$.

Solution. We cannot use the theorem for finding the limit of a quotient in the present case because both the numerator and the denominator of $\frac{x+1}{x}$ are infinitely large quantities and possess no limits. We therefore first transform the fraction as follows, writing

$$\frac{x+1}{x}=1+\frac{1}{x}.$$

We now obtain

$$\lim_{x\to+\infty}\frac{x+1}{x}=\lim_{x\to+\infty}\left(1+\frac{1}{x}\right)=1+\lim_{x\to+\infty}\frac{1}{x}=1+0=1,$$

since $\frac{1}{x}$ is an infinitesimal on indefinite increase of x and its limit is therefore zero.

Example 2. Find $\lim\limits_{x\to+\infty}\dfrac{3x^3-8x}{2x^2-7x^3}$.

Solution. Dividing numerator and denominator by x^3 gives us

$$\frac{3x^3-8x}{2x^2-7x^3}=\frac{\dfrac{3x^3-8x}{x^3}}{\dfrac{2x^2-7x^3}{x^3}}=\frac{3-\dfrac{8}{x^2}}{\dfrac{2}{x}-7}.$$

As $x\to+\infty$, $\frac{8}{x^2}$ and $\frac{2}{x}$ are infinitesimals. If we now apply the theorem regarding the limit of a quotient, we find

$$\lim_{x\to+\infty}\frac{3x^3-8x}{2x^2-7x^3}=\lim_{x\to+\infty}\frac{3-\dfrac{8}{x^2}}{\dfrac{2}{x}-7}=\frac{3}{-7}=-\frac{3}{7}.$$

§ 35. **The ratio of two infinitesimals.** As we shall see later, cases constantly occur in the differential calculus when we have to investigate the ratio of two infinitesimals.

It may be noticed that the theorems regarding the properties of infinitesimals do not include one regarding the ratio of two infinitesimals. This is explained by the fact that, in general, nothing definite can be said about such a ratio, since the nature of its variation depends on the variation of the infinitesimals composing it.

Suppose, for instance, that α is an infinitesimal. Then α^2 and $2\alpha+\alpha^2$ are also infinitesimals.

Now (assuming that α does not take zero values) we have

$\frac{\alpha^2}{\alpha}=\alpha$, which is an infinitesimal;

$\frac{\alpha}{\alpha^2}=\frac{1}{\alpha}$, which is an infinitely large quantity;

$\frac{2\alpha+\alpha^2}{\alpha}=2+\alpha$, which is a bounded quantity.

It may be mentioned that there are more complicated cases of the ratio of two infinitesimals, but the examples given make it perfectly clear that the variation of such a ratio depends on the nature of the variation of the infinitesimals concerned.

EXERCISES

Find the following limits:

1. $\lim\limits_{x\to 1}(x^2-4x+5)$. Ans. 2.

2. $\lim\limits_{x\to c}(x^2-4x+5)$. Ans. c^2-4c+5.

3. $\lim\limits_{x\to -2}\frac{2x^2+1}{x+1}$. Ans. -9.

4. $\lim\limits_{x\to 4}\frac{x^2-2x+3}{x^3-x^2+1}$. Ans. $\frac{11}{49}$.

5. $\lim\limits_{x\to \frac{\pi}{4}}\frac{2x}{\pi-x}$. Ans. $\frac{2}{3}$

6. $\lim\limits_{x\to 0}\dfrac{3x^3+x^2}{x^5+3x^4-2x^2}$. Ans. $-\dfrac{1}{2}$.

7. $\lim\limits_{x\to\frac{3}{2}}\dfrac{4x^2-9}{2x-3}$. Ans. 6.

8. $\lim\limits_{x\to -2}\dfrac{x^4-16}{x+2}$. Ans. -32.

9. $\lim\limits_{x\to 2}\dfrac{x^2+x-6}{x-2}$. Ans. 5.

10. $\lim\limits_{x\to -3}\dfrac{x^4-6x^2-27}{x^3+3x^2+x+3}$. Ans. -7.2.

11. $\lim\limits_{x\to 1}\dfrac{3x^4-4x^3+1}{(x-1)^2}$. Ans. 6.

12. $\lim\limits_{x\to 0}\dfrac{x^2+3x}{x^3}$. Ans. $+\infty$.

13. $\lim\limits_{x\to 1}\dfrac{3x^3-3x^2+2x-2}{(x-1)^3}$. Ans. $+\infty$.

14. $\lim\limits_{x\to 1}\left(\dfrac{1}{1-x}-\dfrac{3}{1-x^3}\right)$. Ans. -1.

15. $\lim\limits_{x\to -2}\left(\dfrac{1}{x+2}-\dfrac{12}{x^3+8}\right)$. Ans. $-\dfrac{1}{2}$.

16. $\lim\limits_{x\to +\infty}\dfrac{x^2+1}{2x^2+2x-1}$. Ans. $\dfrac{1}{2}$

17. $\lim\limits_{x\to -\infty}\dfrac{3x^3-4}{x^2+5x^3}$. Ans. $\dfrac{3}{5}$.

18. $\lim\limits_{x\to +\infty}\dfrac{2x^3-3x^2+4}{5x-x^2-7x^3}$. Ans. $-\dfrac{2}{7}$.

19. $\lim\limits_{x \to +\infty} \dfrac{2x^2-4x+8}{x^3+2x^2-1}$. Ans. 0.

20. $\lim\limits_{x \to +\infty} \dfrac{4x^3-2x+8}{3x^2+1}$. Ans. $+\infty$.

21. $\lim\limits_{x \to -\infty} \dfrac{4x^3-2x+8}{3x^2+1}$. Ans. $-\infty$.

CHAPTER 5

DERIVATIVES

§ 36. Functions. The domain of definition of a function. Notation for functional relationship. Geometric representation of a function. 1. The concept of function is of extreme importance in mathematics in general and occupies a central position in higher mathematics. The reader will already be familiar with the concept from algebra and geometry. It will be sufficient here merely to recall, and define more precisely, the main features that will concern us in higher mathematics. We shall become acquainted at the same time with some new ideas that are needed for studying the differential and integral calculus.

We first of all formulate an accurate definition of function.

The set of all the values that a variable can take under the conditions of a given problem is known as the *domain of variation* of the variable. For instance, if a point moves along the radius r of a circle, its distance from the centre is a variable, the domain of variation being in this case the set of all real numbers from 0 to r. On successively dividing a unit length into 2, 3, 4, ..., n, ... equal parts, the length l of each part is a variable whose domain of variation consists of the infinite set of numbers $\frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots, \frac{1}{n}, \ldots$. In general, the domain of variation will always be a certain set of numbers.

Mathematical analysis is primarily concerned, not with the study of the variation of a single variable on its own, but with a study of the *relationship* between two or more variables when they vary simultaneously.

If there is a connextion between two variables such that the value of one is defined by the value of the other, we say that they are connected by a ***functional relationship***.

According to the conditions of the problem, one variable can usually be assigned arbitrary values from its domain of variation. We know, for instance, that the path s (in metres) traversed by a falling body in time t (seconds) is given in the absence of resistance in the medium by

$$s=\frac{1}{2}gt^2, \tag{1}$$

where $g=9{\cdot}81$ m/sec^2 is the acceleration due to gravity. Equation (1) defines the relationship between t and s.

Suppose we are interested in the distance s traversed by the body in different intervals of time t. We can assign to t any numerical values we like, and find from equation (1) the corresponding values of s that interest us. In other words, (1) defines a connection between s and t such that s, instead of varying arbitrarily, varies only in *relationship* to the variation of t. The variation of s thus differs in character, in the problem as stated, from the variation of t. A difference in the nature of the variation of the variables involves a difference in the names given to them. A quantity which can be assigned arbitrary values from its domain of variation is called an *independent variable* or *argument*. A variable which takes definite numerical values depending on the values assigned to the argument is called a *dependent variable* or *function*.

We now forget the concrete physical significance of the quantities concerned and give a more general definition of function. As we have already mentioned, this is one of the basic concepts of mathematical analysis.

Let two variables x and y be given and assume that the conditions of the problem are such that arbitrary values can be assigned to x from its domain of variation. *If by virtue of some rule a single definite value of y corresponds to each value of x from its domain of variation, y is said to be a function of x.*

The domain of variation of the argument x is termed the *domain of definition of the function y.*

The most commonly encountered domains of definition of functions are the *interval* (or *open interval*) and the *segment* (or *closed interval*).

An interval is defined as the set of all real numbers x lying between two given numbers a and b but excluding a and b themselves. It is written

$$a<x<b. \tag{2}$$

A segment is defined as the set of all real numbers x lying between and including two given numbers a and b. It is written

$$a\leqslant x\leqslant b. \tag{3}$$

We write the interval (2) for brevity as (a, b) and the segment (3) as $[a, b]$.

An interval (a, b) can be represented on the Ox axis as a segment with its ends excluded, which we usually denote by means of arrows with their tips at the points a and b (fig. 39, A), whilst $[a, b]$ is denoted by an ordinary segment of Ox (fig. 39, B).

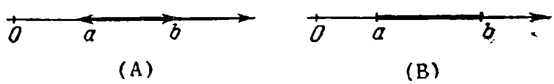


FIG. 39

Examples. (*i*). The domain of definition of the function

$$y=\arcsin x$$

is the segment $[-1, 1]$, i.e. the set of values of x satisfying $-1\leqq x\leqq 1$.

(*ii*) The domain of definition of the function

$$y=\sqrt{4-x^2}$$

is the segment $-2\leqq x\leqq +2$ (for $x<-2$ and $x>2$ the difference $4-x^2<0$, so that $\sqrt{4-x^2}$ is imaginary).

(*iii*). The domain of definition of the function

$$y=\frac{1}{\sqrt{4-x^2}}$$

is the interval $(-2, 2)$ (as in the previous example, $\sqrt{4-x^2}$ is imaginary for $x<-2$ and $x>2$, but now $x=-2$ and $x=2$ are not

permissible values since $4-x^2$ vanishes for them and division by zero is impossible).

When the domain of definition of a function extends over all the real numbers the function is said to be defined in an interval "from minus infinity to plus infinity", which we write symbolically as $(-\infty, +\infty)$. For example, the functions $y=\sin x$ and $y=x^2$ are defined in the interval $(-\infty, +\infty)$. As regards the function $y=\log x$, we can say that it is defined on the positive semi-axis of abscissae, i.e. in the interval $(0, +\infty)$.

2. To indicate that y is a function of the independent variable x we write:

$y=f(x)$, or $y=\varphi(x)$, or $y=F(x)$ and so on. The letters f, φ, F.... symbolize the rule by which the corresponding value of y is obtained for a given value of x. If we are considering several functions simultaneously, characterized by different rules for defining the variation of the function in relation to the variation of the argument, we use different letters in front of the brackets that enclose the argument.

If we want to denote the *particular* value that a function $y=f(x)$ takes for a particular value of x equal to a, we use the symbol $f(a)$. For instance, if

$$y=f(x)=3x^2-8x+2,$$

we have for $x=2$:

$$f(2)=3\cdot 4-8\cdot 2+2=-2;$$

for $x=a$, we have

$$f(a)=3a^2-8a+2;$$

for $x=b-1$, we have

$$f(b-1)=3(b-1)^2-8(b-1)+2=3b^2-14b+13.$$

3. The properties of a function may conveniently be investigated by representing it geometrically, as follows:

Let the domain of definition of the function $y=f(x)$ be the segment $[a, b]$. We shall regard x and y as the co-ordinates of a point $M(x, y)$ on the xOy plane. We plot M (fig. 40) whilst vary-

ing x continuously from the value a to b in such a manner that x takes all the intermediate values between a and b. The segment xM moves along during this process and its end M describes a certain curve. The curve is referred to as the *graph* of the function $f(x)$.

In particular, the graph of $y=f(x)$ may be a straight line; for instance, we know from analytic geometry that the graph of $y=$ $=kx+b$ is a straight line.

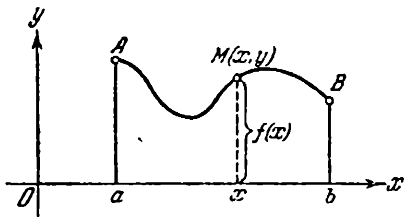


FIG. 40

The graphs of some functions will already have been plotted with the aid of points in courses on algebra and trigonometry, so that we omit examples of this.

We return to the problem of tracing the graphs of functions in Chapter 6, where an investigation of the function with the aid of the differential calculus is used as a basis for tracing its graph.

4. A functional relationship between several (more than two) variables defines a function of several variables.

We take as an example of such a function the volume of a circular cone, which depends on two arguments, i.e. the radius r of the base and the height h:

$$V=\frac{1}{3}\pi r^2 h.$$

The general notation for functions of several variables follows the same principle as for functions of a single variable. For instance, we can express the fact that the volume of a cone is a function of the radius r of the base and the height h by writing

$$V=f(r, h).$$

5. To specify a function means to state the rule or law in accordance with which a given value of the argument x defines a corresponding value of the function y. The definition of function gives no indication of the methods of specifying a function, and in fact these methods can be extremely diverse. Functions in mathematical analysis are mainly specified with the aid of formulae. Elementary algebra and geometry will have made us familiar with plenty of examples of functions specified by formulae.* We shall be concerned with this method of specification in the present course.

In technology and natural science the correspondence between the values of the argument and function is often established with the aid of experiment (observations). For instance, if water is heated to boiling-point in a closed vessel at different pressures p, we can find by observation the temperature T of the boiling-point corresponding to any given pressure. Since the value of T depends on the pressure p, T must be a function of p. Instead of specifying the correspondence between values of the argument p and of the function T by means of a formula, a table is used in the present case (see the table below), which sets out the experimental data. Examples of functions specified by tables may be found in any reference book.

p (in atmospheres)	T (in degrees C)
1	100
1·96	120
3·57	140
6·10	160
15·34	200
39·24	250
84·80	300
163·21	350

It is sometimes possible to use tabulated data to find an approximate formula for the correspondence between argument and func-

* For example, the area S of a square of side x is defined as a function of the independent variable x by the formula $S = x^2$.

tion, in which case the formula is termed *empirical* (i.e. obtained experimentally).

Recording devices are widely used in present-day science and technology which enable functions to be specified by graphs. A device called a barograph, for example, traces a graph showing the variation in atmospheric pressure in the course of a day. Here the graph represents the atmospheric pressure as a function of time.

We shall not go into further details regarding the specification of functions with the aid of tables and graphs since this does not concern us in the present course of mathematical analysis.

6. Let the variables x and y be connected by the equation

$$Ax+By+C=0. \tag{*}$$

If the value of one variable is known, the other cannot be of arbitrary value, since the values of x and y together must satisfy equation (*) (i.e. cause the left-hand side to vanish). The equation thus expresses a functional relationship between x and y.

If we assign the role of argument to one variable, say x, the other variable y becomes a function of x.

We have here an example of specifying a function y with the aid of an equation which is not solved with respect to y. This method of specifying a function is termed *implicit*, the function thus specified being termed an *implicit function*. If we solve equation (*) with respect to y, the function y becomes *explicit*, i.e. it is specified *explicitly*:

$$y=-\frac{Ax+C}{B}.$$

Not every implicit function can be given explicitly. For instance, the equation

$$y+3x-4\sin y=0$$

defines y as a function of x. But this equation cannot be solved for y, so that the present implicit method of specifying y cannot be made explicit.

In general, we can denote an implicit function by writing

$$F(x, y)=0.$$

§ 37. **Increments of argument and function.** Given the function $y=f(x)$, let its argument x vary so as to take the two values x_1 and x_2 in turn. We then call x_2-x_1 the increment of the argument and denote it by Δx_1, i.e.

$$\Delta x_1=x_2-x_1.$$

The Greek letter Δ (delta) is not a factor of x, but merely indicates the operation of subtracting the previous value of the variable, $x=x_1$, from the new value x_2. Thus Δ can no more be separated from x_1 than sin can from x in the expression $\sin x$, or $\log_a$ from x in the expression $\log_a x$.

The difference between the values of the function $f(x)=y$ corresponding to the values x_2 and x_1 of the argument, i.e. $f(x_2)-f(x_1)=y_2-y_1$ is called the increment of the function and is denoted by Δy_1, so that

$$\Delta y_1=f(x_2)-f(x_1)=y_2-y_1.$$

Example. Let the function be $y=x^2-3x$ and let $x_1=2$ and $x_2=2\cdot 3$. Then

$$\Delta x_1=x_2-x_1=2\cdot 3-2=0\cdot 3;$$

$$y_1=2^2-3.2=4-6=-2,$$

$$y_2=(2\cdot 3)^2-3.2\cdot 3=5\cdot 29-6\cdot 9=-1\cdot 61;$$

hence

$$\Delta y_1=y_2-y_1=-1\cdot 61-(-2)=0\cdot 39.$$

If $x_2>x_1$, the increment Δx_1 is positive and if $x_2<x_1$, $\Delta x_1=x_2-x_1$ is negative.

Obviously, the increment of a function can also be positive or negative.

§ 38. **Continuity of functions.** 1. The concept of continuous function plays a fundamental role in mathematical analysis. In particular, the concept is of great importance for our immediate aim of finding the limit of a function.

Suppose that fig. 41 is a representation of the function $y=f(x)$, whilst fig. 42 represents $y=\varphi(x)$. Whereas the first function is represented by a smooth unbroken line, the graph of $y=\varphi(x)$ has a break at $x=x_0$. It is doubtful if anyone will find the statement obscure: "fig. 41 represents a function which is *continuous* for all values of x, whereas fig. 42 represents a function having a *discontinuity* at $x=x_0$ (or at the *point* $x=x_0$ *)".

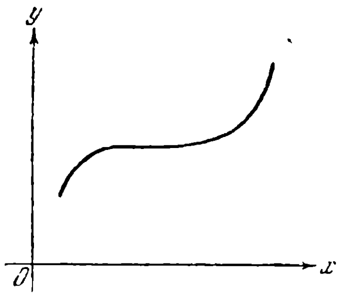


FIG. 41

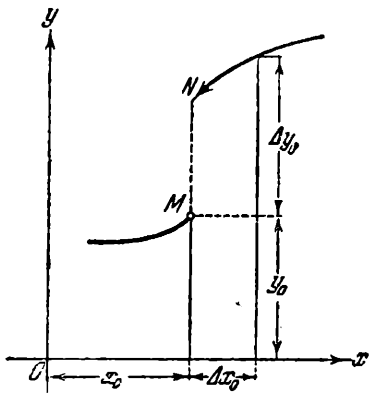


FIG. 42

It will be seen that the discontinuity of $y=\varphi(x)$ at $x=x_0$ amounts to the function changing by a jump as x passes through x_0. The changing of the argument from x_0 to a value x can be represented as the addition of the increment $\Delta x_0=x-x_0$ to x_0. The new value of the function $y=\varphi(x)=\varphi(x_0+\Delta x_0)$ will now differ from the old $y_0=\varphi(x_0)$ by the increment

$$\Delta y_0=\varphi\,(x_0+\Delta x_0)-\varphi\,(x_0).$$

We see from fig. 42 that when Δx_0 tends to zero whilst remaining positive, Δy_0 tends towards a value equal to the length of MN. It is clear that the function will not have a discontinuity at $x=x_0$ if the increment Δy_0 tends to zero.

* Referring to a particular value of the argument as a point derives from the method of representing the value on a number axis.

The above remarks enable us to visualize easily the idea of the continuity of a function at a point. We now proceed to an accurate definition.

We define a *neighbourhood* of x_0 as any interval containing x_0 as an interior point.

DEFINITION. *A function $y=f(x)$ is said to be continuous at the point x_0 if (i) it is defined in a neighbourhood of the point and (ii) the limit of the increment Δy_0 of the function, produced by the increment Δx_0 of the argument x, is zero as $\Delta x_0 \to 0$*, i.e.

$$\lim_{\Delta x_0 \to 0} \Delta y_0 = \lim_{\Delta x_0 \to 0} [f(x_0+\Delta x_0)-f(x_0)]=0. \tag{4}$$

In other words, the function $y=f(x)$ is said to be continuous at x_0 if it is defined in a neighbourhood of x_0 and if its increment Δy_0 is an infinitesimal as $\Delta x_0 \to 0$.

If the continuity condition is not satisfied at x_0, the function is said to have a discontinuity at this point, and x_0 is referred to as a point of discontinuity of the function.

2. If the increment Δy_0 of a function is an infinitesimal as $\Delta x_0 \to 0$, $y=f(x)$ is by definition continuous at x_0. Since

$$\Delta y_0 = f(x_0+\Delta x_0)-f(x_0),$$

it follows that, as $\Delta x_0 \to 0$, the difference between the variable $f(x_0+\Delta x_0)$ and the number $f(x_0)$ is an infinitesimal. Hence we have, by the definition of limit,

$$\lim_{\Delta x_0 \to 0} f(x_0+\Delta x_0)=f(x_0). \tag{5}$$

Let $x_0+\Delta x_0=x$; in this case, when $\Delta x_0 \to 0$, $x \to x_0$. Equation (5) can thus be re-written as

$$\lim_{x \to x_0} f(x)=f(x_0),$$

i.e. *if a function $y=f(x)$, defined in a neighbourhood of the point x_0, is continuous at x_0 if its limit as $x \to x_0$ is equal to the value of the function at $x=x_0$.*

Conversely, suppose we know that

$$\lim_{x \to x_0} f(x)=f(x_0). \tag{6}$$

It follows that

$$f(x)-f(x_0)$$

is an infinitesimal as $x \to x_0$. But this difference is the increment Δy_0 of the function corresponding to the increment Δx_0 of the argument. When $x \to x_0$, $\Delta x_0 \to 0$. Thus if equation (6) is valid, we have $\lim\limits_{\Delta x_0 \to 0} \Delta y_0 = 0$, i.e. the function $y=f(x)$ is continuous at the point x_0.

These last propositions show that the above definition of continuity at a point can be replaced by the following:

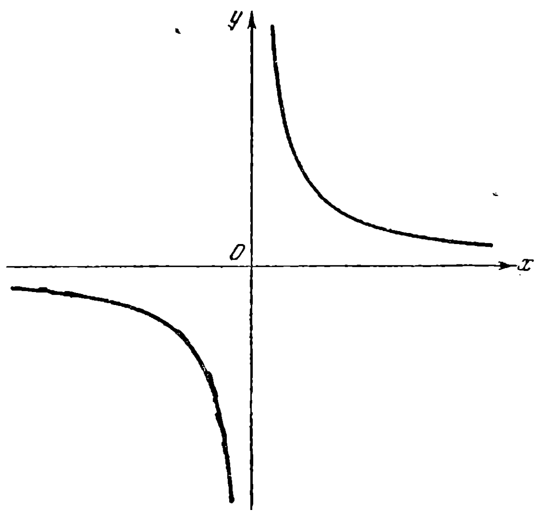


FIG. 43

A function $y=f(x)$ is said to be continuous at the point x_0 if it is defined in a certain neighbourhood of the point and if

$$\lim_{x \to x_0} f(x) = f(x_0). \tag{7}$$

3. We have already encountered discontinuous as well as continuous functions. For instance, the function $y=x^2$ is defined for all real numbers and hence in any neighbourhood of any given point $x=c$; we can show that it is continuous at $x=c$. For $x=c$,

we have $y=c^2$. On giving the value $x=c$ the increment Δx, we find

$$\Delta y=(c+\Delta x)^2-c^2=2c\,\Delta x+(\Delta x)^2 .$$

As $\Delta x\to 0$, $2c\Delta x\to 0$ and $(\Delta x)^2\to 0$. Hence Δy is an infinitesimal as $\Delta x\to 0$, or in other words, $y=x^2$ is continuous at $x=c$.

The functions $y=\dfrac{1}{x}$ and $y=\dfrac{1}{x^2}$ are defined for all real numbers except $x=0$, at which point the expressions become meaningless (division by zero being impossible). Condition (7) for the con-

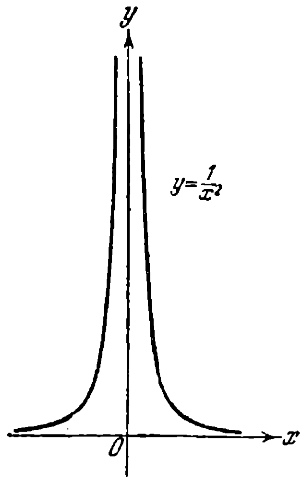


FIG. 44

tinuity of $f(x)$ at the point $x=x_0$ requires that $f(x)$ be defined in a certain neighbourhood of x_0, and therefore at x_0 itself. Since the present functions are not defined at $x=0$, they are discontinuous at this point. They become in fact infinitely large magnitudes as $x\to 0$. Their graphs are illustrated in figs. 43 and 44 respectively. It will be observed that each graph consists of two separate branches; this serves as a visual illustration of the fact that $x=0$ is a point of discontinuity for the functions.

4. *If a function is continuous at every point of a segment* $[a,b]$ *or of an interval* (a, b), *we say that it is continuous in the segment* $[a, b]$ *or in the interval* (a, b).

Thus $y=x^2$ is continuous in the interval $(-\infty, +\infty)$, whilst $y=\frac{1}{x}$ and $y=\frac{1}{x^2}$ are continuous in any segment not including the point $x=0$.

The function $\tan x$ is continuous for all values of x except $x=\frac{1}{2}\pi+k\pi$ (where $k=0, \pm 1, \pm 2, \ldots$) as the continuity condition is not fulfilled at these points (since $\tan x$ is not defined for $x=\frac{2}{2}\pi+k\pi$). Similarly, $\cot x$ has discontinuities at the points $x=k\pi$ (where $k=0, \pm 1, \pm 2, \ldots$). For all other values of x, $\cot x$ is continuous.

Complete courses of mathematical analysis treat in detail the continuity of the functions to be analysed. We merely mention here that all the functions to be encountered below are continuous everywhere except possibly at certain specific points.

5. If $f(x)$ is continuous at the point c we have

$$\lim_{x\to c} f(x)=f(c).$$

Hence, to find the limit of $f(x)$ as $x\to c$, we only have to work out the value of $f(x)$ for $x=c$, the required limit being in fact the number thus obtained.

We can find in this way, for example:

$$\lim_{x\to\frac{\pi}{2}} \sin x=\sin\frac{\pi}{2}=1;$$

$$\lim_{x\to\frac{\pi}{4}} \tan x=\tan\frac{\pi}{4}=1;$$

$$\lim_{x\to\frac{1}{2}\pi} \log\sin x=\log\sin\frac{1}{2}\pi=\log 1=0$$

and so on.

Moreover, this simplicity in finding the limits of continuous functions enables the limits to be found fairly easily in a number

of cases of functions that are not continuous at the point concerned.

Suppose, for instance, that we want to find

$$\lim_{x \to \frac{\pi}{2}} \frac{\sin 2x}{\cos x}.$$

At $x = \frac{1}{2}\pi$ the function $\frac{\sin 2x}{\cos x}$ has a discontinuity, for $\cos\frac{1}{2}\pi = 0$, and since division by zero is impossible, the fraction $\frac{\sin 2x}{\cos x}$ becomes meaningless as $x = \frac{1}{2}\pi$. In other words, our function is not defined at $x = \frac{1}{2}\pi$ and therefore has a discontinuity at this point. But we leave out of account the value $x = \frac{1}{2}\pi$ as $x \to \frac{1}{2}\pi$ (see § 31). As $x \to \frac{1}{2}\pi$, $\cos x \neq 0$ and we can cancel $\frac{\sin 2x}{\cos x}$ by $\cos x$ to give

$$\frac{\sin 2x}{\cos x} = \frac{2\sin x \cos x}{\cos x} = 2\sin x.$$

It follows from this that $\frac{\sin 2x}{\cos x}$ and $2\sin x$ vary in the same manner as $x \to \frac{1}{2}\pi$. The function $2\sin x$ is continuous at $x = \frac{1}{2}\pi$ and therefore, $\lim\limits_{x \to \frac{1}{2}\pi} (2\sin x) = 2\sin\frac{1}{2}\pi = 2$. Hence we also have $\lim\limits_{x \to \frac{1}{2}\pi} \frac{\sin 2x}{\cos x} = 2$. Similarly, we find that

$$\lim_{x \to 4} \frac{\sqrt{x+5}-3}{x-4} = \lim_{x \to 4} \frac{(\sqrt{x+5}-3)(\sqrt{x+5}+3)}{(x-4)(\sqrt{x+5}+3)} =$$

$$= \lim_{x \to 4} \frac{x+5-9}{(x-4)(\sqrt{x+5}+3)} = \lim_{x \to 4} \frac{1}{\sqrt{x+5}+3} = \frac{1}{6}.$$

§ 39. **Uniform motion and its velocity. Rate of change of a linear function.** 1. If the ratio of the distance traversed by a particle in a given interval of time to this interval of time is constant, the motion of the particle is said to be *uniform*. The constant ratio gives the distance traversed by the particle in unit time (per second, minute, hour) and is termed the *velocity* of the uniform motion.

Suppose that our uniformly moving particle has already traversed a distance s_0 at the initial instant, which we take equal to zero. Let t be any interval of time measured from 0. Let s denote the length of path traversed by the particle at the instant t. Since s is measured from the same origin as s_0, the distance traversed in time t is given by $s-s_0$. By definition of uniform motion, the ratio

$$\frac{s-s_0}{t}=v$$

is constant and represents the velocity of the particle. We obtain from this equation:

$$s=vt+s_0 \tag{8}$$

Equation (8) is called the law of uniform motion, and will be seen to consist of a function of the first degree in t. *Thus the law of uniform motion is given by a function of the first degree in t.* The (constant) coefficient v is the velocity of the uniform motion.

Conversely, it is easily shown that the function

$$s=kt+b \tag{9}$$

of the first degree in t (k and b are constants) represents a uniform motion. For, given any instant t, the path traversed at t is given by (9). Let t_1 now be some other instant, and let s_1 be the path traversed at t_1. We have by (9)

$$s_1=kt_1+b.$$

Thus the distance traversed in time t_1-t is

$$s_1-s=(kt_1+b)-(kt+b),$$

or

$$s_1 - s = k(t_1 - t),$$

whence

$$\frac{s_1 - s}{t_1 - t} = k. \tag{10}$$

Since k is constant, we conclude that the motion defined by $s = kt + b$ is characterized by the constancy of the ratio of the distance traversed in any given interval of time to this interval of time. But, in fact, this is the definition of uniform motion.

Now $t_1 - t$ is the increment Δt of time, whilst $s_1 - s = \Delta s$ is the increment of the distance (see § 37). We can thus say from equation (10) that the velocity v of uniform motion is the ratio of the increment Δs of the distance s to the increment Δt of time t that produces it, i.e.

$$v = \frac{\Delta s}{\Delta t}.$$

2. The variation in the length l of a spring is given in terms of the load P by the law

$$l = kP + l_0, \tag{11}$$

where k and l_0 are constants (the familiar Hooke's law of mechanics). We obtain, as in the case of uniform motion:

$$\frac{\Delta l}{\Delta P} = k.$$

This ratio gives the elongation of the spring produced by unit load. It therefore shows how fast the spring length changes as the load varies. We can thus call $\frac{\Delta l}{\Delta P}$ the velocity (or rate) of change of the spring length relative to the change in load.

From equation (11), the spring length l is a function of the first degree in the load P. The process of varying the length in accordance with the load is thus a uniform process. We see once again that the velocity of a uniform process is given by the ratio of the increment Δl of the function l to the increment ΔP of the argument P producing it.

3. We take the function

$$y = kx + b. \tag{12}$$

If we regard x and y as mathematical quantities i.e. as separated from any physical significance, the first degree function (12) will simply express the variation of the variable y in accordance with the variation of the quantity x.

The ratio $\dfrac{\Delta y}{\Delta x} = k$ now defines the variation in the function y occurring with unit change in the argument x. It seems natural, by analogy with the velocity of uniform motion and the rate of change of the length of a spring, to term $\dfrac{\Delta v}{\Delta x}$ *the rate of change of the function y with respect to the argument x.*

We know from analytic geometry that the graph of the first degree function (12) is a straight line. For this reason, (12) is also termed a *linear* function. The coefficient k is known is the slope of the straight line.

We shall also call this coefficient the slope of the linear function. The relationship

$$\frac{\Delta y}{\Delta x} = k$$

shows that the *rate of change of the linear function*

$$y = kx + b$$

with respect to the argument x is a constant equal to the slope k of the function.

The idea thus established of the rate of change of a linear function is in full agreement with the geometrical picture, for, the greater the slope, the steeper the rise of the straight line representing function (12), and the faster the growth of the ordinate of a point of the straight line corresponding to a growth of the abscissa x.

§ 40. Non-uniform motion and its velocity. 1. The previous section has established the idea of the rate of change of any uniform

variation. The majority of the variations found in nature and subject to scientific inquiry are, however, non-uniform. We need only think of a heavy particle falling in space under the action of gravity. We know from physics that a particle falls in space according to the law expressed by

$$s = \frac{g}{2} t^2, \tag{13}$$

where t is time, measured from the start of the fall, s is the distance traversed in time t, and g is the acceleration due to gravity, equal to 9·81 m/sec². The motion here is non-uniform since the law concerning it is expressed by a function of the second degree in t, whereas, as we have seen, the law for any uniform motion is given by a function of the first degree in t.

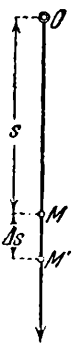


FIG. 45

We now consider the possibility of defining the rate of change of a non-uniform variation. We shall start by considering the velocity of a non-uniform motion, and return to our example of a body falling in space.

We have seen that the velocity of uniform motion can be defined in magnitude as the ratio $\frac{\Delta s}{\Delta t}$. We shall start fiom the same sort of ratio when considering the velocity of non-uniform motion.

Let us take a definite instant t and let the particle be at the position M (fig. 45) at this instant. The distance $s=OM$ traversed by the particle after time t is given by equation (13):

$$s=\frac{g}{2}t^2.$$

Suppose that the time is increased by the amount Δt, i.e. t is given the increment Δt. Let the particle be at M' at the instant $t+\Delta t$. Let Δs denote the path increment MM' corresponding to the time increment Δt. On replacing t in equation (13) by $t+\Delta t$, we find for the new path OM'

$$s+\Delta s=\frac{g}{2}(t+\Delta t)^2.$$

On subtracting equation (13) from this, we obtain Δs:

$$\begin{array}{r l} s+\Delta s= & \frac{g}{2}(t+\Delta t)^2 \\ s= & \frac{g}{2}t^2 \\ \hline \Delta s= & \frac{g}{2}[t^2+2t\cdot\Delta t+(\Delta t)^2-t^2] \end{array}$$

or

$$\Delta s=\frac{g}{2}[2t\cdot\Delta t+(\Delta t)^2].$$

We divide Δs by Δt and obtain an expression giving $\frac{\Delta s}{\Delta t}$:

$$\frac{\Delta s}{\Delta t}=\frac{g}{2}\cdot\frac{2t\cdot\Delta t+(\Delta t)^2}{\Delta t},$$

or, on cancelling Δt,

$$\frac{\Delta s}{\Delta t}=\frac{g}{2}(2t+\Delta t). \tag{14}$$

The ratio $\frac{\Delta s}{\Delta t}$ defining the velocity of motion, is constant when the motion is uniform, i.e. it remains the same for any t or Δt.

In other words, we can take any instant t and any increment Δt in order to find the velocity of uniform motion.

The situation is different in the case of non-uniform motion. In our example of a falling body, we see from equation (14) that $\frac{\Delta s}{\Delta t}$ depends on both t and Δt. For a given value of Δt, we get different values of $\frac{\Delta s}{\Delta t}$ corresponding to different instants. For instance, for $\Delta t = 0{\cdot}1$ sec, we get a value corresponding to $t = 1$ sec of

$$\frac{\Delta s}{\Delta t} = \frac{1}{2}\, g\,(2.\ 1 + 0{\cdot}1) = 2{\cdot}1.\ \frac{1}{2}\, g = 1{\cdot}05\, g,$$

and corresponding to $t = 3$ sec

$$\frac{\Delta s}{\Delta t} = \frac{1}{2}\, g\ (2.3 + 0{\cdot}1) = 6{\cdot}1.\ \frac{1}{2}\, g = 3{\cdot}05\, g,$$

and so on. The velocity of non-uniform motion can thus only be thought of in reference to a definite instant. We cannot speak of the velocity in general of a falling body, but only of the velocity with which the body is falling at a given instant, i.e. of an instantaneous velocity of falling.

But how shall we define instantaneous velocity?

Equation (14) further shows us that, with t fixed, $\frac{\Delta s}{\Delta t}$ depends on Δt. Thus, suppose we take $t = 3$, we get for $\Delta t = 0{\cdot}5$:

$$\frac{\Delta s}{\Delta t} = \frac{1}{2} g\ (2.3 + 0{\cdot}5) = 6{\cdot}5.\ \frac{1}{2} g = 3{\cdot}25\, g,$$

whilst for $\Delta t = 0{\cdot}1$ we have

$$\frac{\Delta s}{\Delta t} = \frac{1}{2} g\ (2.3 + 0{\cdot}1) = 6{\cdot}1.\ \frac{1}{2} g = 3{\cdot}05\, g,$$

and so on.

Let us suppose for the moment that the particle moves uniformly from the position M to M' (fig. 45). Its velocity of falling

would then equal to $\frac{\Delta s}{\Delta t}=\frac{1}{2}g(2t+\Delta t)$ throughout the interval Δt and would remain constant. Just as we might say, for instance, that a train travelling from London to Carlisle has an average velocity of 50 m.p.h. over the section from Crewe to Carnforth, so we refer here to the constant velocity $\frac{\Delta s}{\Delta t}$ as the *average velocity* of falling of the particle over the section MM'. Let v_{av} denote this average velocity.

As we have seen, the average velocity varies as Δt varies, and it is perfectly clear that it gives a better indication of the state of the falling particle at the instant t, the smaller we take the interval Δt succeeding the instant t. It seems natural to conclude from this that *the instantaneous velocity of a falling particle at the instant t is to be defined as the limit of the average velocity* v_{av} *as* $\Delta t \to 0$, i.e.,

$$v=\lim_{\Delta t\to 0} v_{av}.$$

We thus obtain

$$v=\lim_{\Delta t\to 0}\frac{\Delta s}{\Delta t}=\lim_{\Delta t\to 0}\left[\frac{g}{2}(2t+\Delta t)\right]=gt.$$

2. We turn to the general case of non-uniform motion and let s denote the distance traversed by a particle after a time t. Since a definite value of distance traversed s corresponds to each value of t, s is a function of t, i.e.

$$s=f(t).$$

To find the instantaneous velocity of the particle at a given instant t, we first find the average velocity of the motion over the interval from t to $t+\Delta t$. The particle has traversed a distance $s=f(t)$ after time t and after time $t+\Delta t$ the distance will be $s+\Delta s=f(t+\Delta t)$. The distance Δs traversed in the interval Δt is therefore given by

$$\Delta s=f(t+\Delta t)-f(t).$$

We obtain the average velocity v_{av} over the interval Δt by dividing Δs by Δt:

$$v_{av}=\frac{\Delta s}{\Delta t}=\frac{f(t+\Delta t)-f(t)}{\Delta t}.$$

On now passing to the limit as $\Delta t\to 0$, we obtain the instantaneous (or true) velocity of the particle at the instant t:

$$v=\lim_{\Delta t\to 0} v_{av}=\lim_{\Delta t\to 0}\frac{\Delta s}{\Delta t}=\lim_{\Delta t\to 0}\frac{f(t+\Delta t)-f(t)}{\Delta t}.$$

Thus *the instantaneous velocity v of a particle at the instant t is defined as the limit to which the average velocity v_{av} over the interval Δt tends as $\Delta t\to 0$.*

3. *Note* 1. The definition of the velocity of non-uniform motion embraces the velocity of uniform motion as a particular case. For, since the velocity $\frac{\Delta s}{\Delta t}$ of uniform motion is constant, and the limit of a constant is equal to the constant itself, we have for the case of uniform motion

$$\lim_{\Delta t\to 0}\frac{\Delta s}{\Delta t}=\frac{\Delta s}{\Delta t}.$$

Note 2. In all our arguments so far we have regarded the instant $t+\Delta t$ as coming later than the instant t, i.e. we have always assumed that Δt is positive. It makes no difference, however, if $t+\Delta t$ is looked on as coming earlier than t. For we defined the velocity v as the limit of the average velocity v_{av} as $\Delta t\to 0$, and the definition of limit requires that v_{av} tends to a definite number independently of the manner in which Δt tends to 0.

§ 41. The rate of change of a function (the basic problem leading to the concept of derivative). 1. We established in § 39 the idea of the rate of uniform change of a single variable depending on the change of some other variable. We now define a rate of change for any non-uniform variation.

We take the function $y=f(x)$ and regard x and y as mathematical quantities, i.e. we do not attribute a physical significance to them.

In this case $y=f(x)$ simply expresses the variation of y in accordance with the variation of x. Let a given value of the argument x be assigned the increment Δx, and let the corresponding increment of y be Δy. The ratio $\frac{\Delta y}{\Delta x}$ shows how much faster or slower "on the average" the variation of y is than that of x, when x changes by the amount Δx. Hence we may term $\frac{\Delta y}{\Delta x}$ the average rate of change (v_{av}) of y with respect to x when x changes by Δx. It now seems natural to define the rate of change (v) of y for a given value of x as the limit of $\frac{\Delta y}{\Delta x}$ as $\Delta x \to 0$, i.e.

$$v = \lim_{\Delta x \to 0} v_{av} = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}.$$

The rate of change of the function y for a given value of the argument x is defined as the limit of the ratio of the increment Δy of the function to the increment Δx of the argument as $\Delta x \to 0$.

2. The above generalized rate of change of any type of variation is applicable in various departments of science. We have seen, for example, that if x is taken to denote time, and y the distance traversed by a particle in the time x, $\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}$ defines the velocity of the particle at a given instant x.

Obviously, both the average velocity $\frac{\Delta y}{\Delta x}$ and the instantaneous velocity $\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x}$ of the particle are expressed by denominate, not by abstract, numbers. The unit of velocity used here is the velocity of the uniform motion in which the distance y changes by unit distance Δy when the time x changes by unit time Δx. For instance, if y changes by $\Delta y = 1$ cm when x changes by $\Delta x = 1$ sec, the unit of velocity is $\frac{\Delta y}{\Delta x} = 1$ cm/sec; if $\Delta y = 1$ m for $\Delta x = 1$ min, the unit of velocity is $\frac{\Delta y}{\Delta x} = 1$ m/min, and so on.

The position is similar whenever x and y have some sort of physical significance, i.e. when it is a question of the velocity of some real physical process.

The definition of velocity or rate of change for any process forms the fundamental step leading to the concept of derivative. This concept will be formulated precisely in the next section. For the moment, we consider a few particular examples of our basic problem:

1) When a rigid body rotates about an axis, its angle of rotation φ is clearly a function of time t. If the body rotates through the angle $\Delta\varphi$ in the interval from t to $t+\Delta t$, the ratio $\frac{\Delta\varphi}{\Delta t}$ gives the average angular velocity in the interval Δt, whilst $\lim\limits_{\Delta t\to 0}\frac{\Delta\varphi}{\Delta t}$ gives the angular velocity at the instant t.

2) Let Q denote the amount of a substance taking part in a chemical reaction at the instant t. The ratio $\frac{\Delta Q}{\Delta t}$ now gives the average velocity of the reaction for the interval from t to $t+\Delta t$, whilst $\lim\limits_{\Delta t\to 0}\frac{\Delta Q}{\Delta t}$ gives the reaction velocity at the instant t.

3) Let W be the amount of heat (in calories) required to raise the temperature of a body from 0 to θ (degrees C). It is clear that W is a function of θ, i.e. $W=f(\theta)$. Let ΔW be the increment of W corresponding to the increment $\Delta\theta$ of θ. The ratio $\frac{\Delta W}{\Delta\theta}$ gives us the mean increment in the amount of heat reckoned per unit increment in the temperature θ. This mean rate of change of the amount of heat is termed the average heat capacity of the body on heating from θ° to $(\theta+\Delta\theta)^\circ$. The rate of change of the amount of heat at a given temperature θ, i.e. $\lim\limits_{\Delta\theta\to 0}\frac{\Delta W}{\Delta\theta}$, is termed the heat capacity of the body at the given temperature θ.

4) Let Q be the amount of electric charge (in coulombs) flowing through a circuit at time t. The current I is defined as

$$\lim_{\Delta t\to 0}\frac{\Delta Q}{\Delta t}.$$

This limit gives the charge flowing through the circuit per unit time. The current is thus the rate of flow of charge through the conductor.

§ 42. Derivatives. General method of finding derivatives. 1. We have seen above that evaluation of the limit of a ratio of the form

$$\frac{\Delta y}{\Delta x}=\frac{f(x+\Delta x)-f(x)}{\Delta x}$$

as $\Delta x\to 0$ is closely related to a number of fundamental concepts in widely differing departments of science. This limit is therefore given special attention in mathematical analysis and receives a special name. It is called *the derivative of the function $y=f(x)$ with respect to the independent variable x*. The derivative is accurately defined as follows:

The limit as $\Delta x\to 0$ of the ratio of the increment Δy of the function $y=f(x)$ to the increment Δx of the independent variable x that produces Δy is called the derivative of y with respect to x at the given value of x (at the given point x).

We denote the derivative by the symbols y' or $f'(x)$. Thus

$$y'=f'(x)=\lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x}=\lim_{\Delta x\to 0}\frac{f(x+\Delta x)-f(x)}{\Delta x}.$$

Using the definition just given, the results obtained in the previous section may be formulated as follows:

1) *The rate of change of the function $y=f(x)$ at a given value of x is the derivative of y with respect to x at the given point;*

2) *the instantaneous velocity is the derivative of the distance s traversed with respect to time;*

3) *the angular velocity of a body rotating about an axis is the derivative of the angle φ of rotation with respect to time t;*

4) *the velocity of a chemical reaction is the derivative of the amount Q of substance taking part in the reaction with respect to time t;*

5) *the heat capacity of a body is the derivative of the amount of heat W absorbed by the body with respect to the temperature θ;*

6) *the current is the derivative of the charge flowing Q with respect to time t.*

The differential calculus is primarily concerned with evaluating derivatives and studying their properties and applications.

2. The process of evaluating a derivative is termed *differentiation.* Thus to speak of "differentiating a function" is the same as saying "evaluating (or finding) the derivative of the function".

By the definition of derivative, to differentiate a function y of x we have to carry out the following operations (*general rule for evaluation of derivatives*):

1) find the value of the function y corresponding to the given value of the argument x;

2) give the given value of x the increment Δx and work out the new value $y+\Delta y$ of the function;

3) subtract the previous value of the function from the new value and thus obtain he increment Δy;

4) form the ratio $\frac{\Delta y}{\Delta x}$, i.e. divide the value obtained for Δy by Δx;

5) find the limit of the ratio $\frac{\Delta y}{\Delta x}$ as $\Delta x \to 0$ to give the required derivative.

We take a number of examples illustrating the general rule for differentiation. We write $y_{x=c}$ for the value of the function y corresponding to the argument $x=c$.

Example 1. Find the derivative of $y=x^2$ at $x=3$ (at the point $x=3$).

Solution.

1) $y_{x=3}=3^2=9;$

2) $y_{x=3}+\Delta y=(3+\Delta x)^2;$

3)
$$\begin{array}{r} y_{x=3}+\Delta y=9+6\Delta x+(\Delta x)^2 \\ - \\ y_{x=3} \qquad =9 \qquad\qquad\quad \\ \hline \Delta y=6\Delta x+(\Delta x)^2; \end{array}$$

4) $\dfrac{\Delta y}{\Delta x}=\dfrac{6\Delta x+(\Delta x)^2}{\Delta x}=6+\Delta x$ *;

5) $y'_{x=3}=\lim\limits_{\Delta x\to 0}(6+\Delta x)=6.$

Example 2. Find the derivative of $y=x^3$ at any point x (for any value of the argument x).

Solution. Since the value of the function y for any value of x is given by the equation $y=x^3$, the first operation reduces simply to writing:

1) $y \quad =x^3;$

2) $y+\Delta y=(x+\Delta x)^3;$

3) $$\begin{array}{rl} & y+\Delta y=x^3+3x^2\,\Delta x+3x\,(\Delta x)^2+(\Delta x)^3 \\ - & y \quad =x^3 \\ \hline & \Delta y=3x^2\,\Delta x+3x\,(\Delta x)^2+(\Delta x)^3; \end{array}$$

4) $\dfrac{\Delta y}{\Delta x}=\dfrac{3x^2\,\Delta x+3x\,(\Delta x)^2+(\Delta x)^3}{\Delta x}=3x^2+3x\cdot\Delta x+(\Delta x)^2$ *;

5) $y'=\lim\limits_{\Delta x\to 0}\dfrac{\Delta y}{\Delta x}=\lim\limits_{\Delta x\to 0}[3x^2+3x\cdot\Delta x+(\Delta x)^2]=3x^2.$

Example 3. Find the derivative of $y=3x^2+5$ at any point x.
Solution.

1) $y=3x^2+5;$

2) $y+\Delta y=3(x+\Delta x)^2+5;$

3) $$\begin{array}{rl} & y+\Delta y=3x^2+6x\cdot\Delta x+3\,(\Delta x)^2+5 \\ - & y \quad =3x^2 \qquad\qquad\qquad\quad +5 \\ \hline & \Delta y=6x\cdot\Delta x+3\,(\Delta x)^2; \end{array}$$

---

* Dividing through by Δx is quite justifiable here, since the limit sought in the fifth operation is with $\Delta x\to 0$, when, as we know, the value $\Delta x=0$ is excluded (cf. § 31).

4) $\frac{\Delta y}{\Delta x} = 6x + 3 \cdot \Delta x;$

5) $y' = \lim\limits_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim\limits_{\Delta x \to 0} (6x + 3\Delta x) = 6x.$

Example 4. Differentiate $y = \frac{c}{x^2}$.

Solution.

1) $y = \frac{c}{x^2};$

2) $y + \Delta y = \frac{c}{(x + \Delta x^2)};$

3) $$\begin{array}{rl} y + \Delta y = & \frac{c}{(x+\Delta x)^2}; \\ - \quad y \quad = & \frac{c}{x^2} \\ \hline \end{array}$$

$$\Delta y = \frac{c}{(x+\Delta x)^2} - \frac{c}{x^2} = -\frac{c \cdot \Delta x (2x + \Delta x)}{x^2 (x + \Delta x)^2}$$

4) $\frac{\Delta y}{\Delta x} = -c \cdot \frac{2x + \Delta x}{x^2 (x + \Delta x)^2};$

5) $y' = \lim\limits_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = -c \cdot \frac{2x}{x^4} = -\frac{2c}{x^3}.$

Example 5. Differentiate $y = kx + b$.

Solution.

1) $y = kx + b;$

2) $y + \Delta y = k(x + \Delta x) + b;$

3) $$\begin{array}{rlll} y + \Delta y & = kx & + k \cdot \Delta x & + b \\ - \quad y & = kx & & + b \\ \hline \end{array}$$

$$\Delta y = k \cdot \Delta x;$$

4) $\dfrac{\Delta y}{\Delta x}=k$;

5) $y'=\lim\limits_{\Delta x \to 0} \dfrac{\Delta y}{\Delta x}=k,$

since k is a constant.

We established in § 39 that the rate of change of a linear function is given by the coefficient k and is therefore constant for any value of x. In § 41 we interpreted the derivative of any function as its rate of change for the given value of x. It will be seen from this that the result now obtained is in full agreement with the conclusion reached in § 39.

Example 6. Find $f'(2)$ and $f'(-3)$ for $f(x)=3x^2+5$ (i.e. find the derivative of the function at the points $x=2$ and $x=-3$).

Solution. We found in example 3 an expression for the derivative of the given function for any x, i.e.

$$y'=f'(x)=6x .$$

The symbol $f'(2)$ denotes the derivative of $f(x)$ at $x=2$. Hence

$$f'(2)=6.2=12 \quad \text{and} \quad f'(-3)=6.-3=-18 .$$

Note. We found the derivative of $f(x)=3x^2+5$ at $x=2$ and $x=-3$ in the last example by substituting the numbers 2 and -3 in turn for x in the expression for the derivative for *any* value of x. We obtain different values of the derivative $f'(x)=6x$ corresponding to different values of x; in other words, the derivative $f'(x)=6x$ of $f(x)=3x^2+5$ is itself a function of x.

Thus the search for a general expression for the derivative $f'(x)$ of a given $f(x)$ at any x leads to forming from $f(x)$ a new function $f'(x)$ of the same argument. We sometimes call $f'(x)$ the *derived function* to distinguish it from the derivative of $f(x)$ at a given point, i.e. for a given numerical value of the argument.

We give a strict definition of derived function.

DEFINITION. *Let the function $f(x)$ be defined in the segment $[a, b]$ (or in the interval (a, b)) and have a derivative at each point of this segment (or interval). The function $f'(x)$ whose value at every*

given point $x=x_0$ is equal to the derivative $f'(x_0)$ at $x=x_0$ of the given $f(x)$ is termed the derived function of $f(x)$.

It follows from this that the derivative at a point is a particular value of the derived function. Finding the derivative for *any value of x* is the same as finding the derived function. We found in examples 2, 3, 4, 5 above the derived functions of

$$y=x^3, \quad y=3x^2+5, \quad y=\frac{c}{x^2}, \quad y=kx+b.$$

It must be mentioned that it is always clear in practice whether we are concerned with the derivative at a point or with the derived function. Hence, whatever the problem, we usually refer simply to the derivative of a function.

§ 43. The slope of a curve. Tangent to a curve. 1. We have already discussed the slope k of the straight line $y=kx+b$. By using the concept of derivative, we can define a slope for any type of curve. Since the slope of the curve will be different at different points, however, the definition must relate to a given point of the curve and not to the curve as a whole.

We mean by the slope of a curve at a given point the slope of the tangent to the curve at that point.

Thus, to find the slope of the curve we have to know the slope of the tangent to the curve in terms of the equation $y=f(x)$ of the curve and the co-ordinates (x, y) of our given point.

We must now define the tangent to a curve.

2. The tangent to a circle is defined in elementary geometry as a straight line that cuts the circle in a single point. We cannot use this as a definition for curves in general. For suppose we tried to apply this definition to the parabola $y=x^2$ (see fig. 48), both the co-ordinate axes would satisfy the definition at the origin 0. Yet it is clear that only the Ox axis is tangential to the parabola at the origin. Since our simple definition cannot be applied to curves in general, we have to look for a more general definition.

We define the tangent to a given curve at the point M_0 by proceeding as follows: we take another point M of the curve (fig. 46)

and draw the chord MM_0. As the point M moves along the curve, the chord M_0M rotates about M_0.

The tangent to the curve at the point M_0 is defined as the limiting position M_0T of the chord M_0M as the point M tends to M_0 along the curve.

If the curve is given by the equation $y=f(x)$, we can draw the tangent to it at the point M_0 (x_0, y_0) provided we know the slope

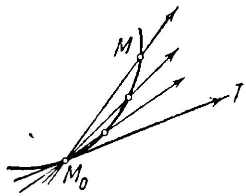


FIG. 46

of the tangent. We consider how to find the slope k of the tangent at M_0.

On giving the abscissa x_0 of M_0 the increment Δx, we pass from M_0 to the point M with abscissa $x_0+\Delta x$ and ordinate y_0+ $+\Delta y=f(x_0+\Delta x)$ (fig. 47). We find the slope $\tan\alpha$ of the chord

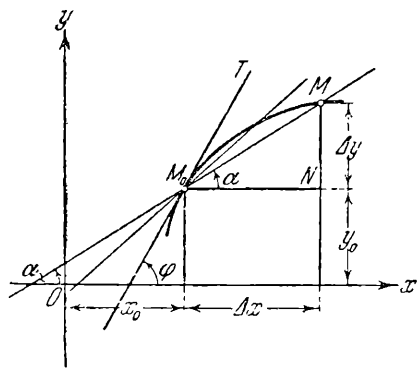


FIG. 47

M_0M from the right-angled triangle M_0NM. Here, M_0N is equal to the increment Δx of x_0 and NM is equal to the corresponding increment of the ordinate y. We have

$$NM = \Delta y = f(x_0 + \Delta x) - f(x_0) .$$

Thus

$$\tan \alpha = \frac{\Delta y}{\Delta x} = \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} .$$

When M tends to M_0 along the curve, the increment $\Delta x \to 0$. To find the slope of the tangent we must therefore find the limit of $\frac{\Delta y}{\Delta x}$ as $\Delta x \to 0$. On writing φ for the angle of inclination of the tangent to Ox, we arrive at the result

$$k = \tan \varphi = \lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = \lim_{\Delta x \to 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} .$$

Since $\lim\limits_{\Delta x \to 0} \frac{\Delta y}{\Delta x}$ is the derivative $f'(x_0)$ of the function $y = f(x)$ at $x = x_0$, we conclude that *the slope k of the tangent to the curve $y = f(x)$ at the point M_0 (x_0 , y_0) is equal to the value of the derivative $f'(x)$ of $y = f(x)$ at $x = x_0$*, i.e. $k = f'(x_0)$.

In other words, *the slope k of the tangent is the derivative at x_0 of the ordinate $y = f(x)$ with respect to the abscissa x.*

We have now solved the problem of finding the slope of a curve at a given point, since the slope of a curve is by definition the slope of the tangent to the curve at the given point.

Example 1. Find the slope of the curve $y = x^3$ at $(2 , 8)$.

Solution. We found in § 42, example 2, that the derived function of $y = x^3$ is $y' = 3x^2$. To find the slope at the point whose abscissa is 2, we substitute 2 for x in the expression for the derived function, and thus obtain

$$k = y' = 3 \cdot 2^2 = 12 .$$

Example 2. Find the equation of the tangent to the parabola $y = x^2$, (*a*) at $M_0\left(\frac{1}{2} , \frac{1}{4}\right)$; (*b*) at the origin.

Solution. (*a*) Since the tangent is a straight line passing through the point $\left(\frac{1}{2} , \frac{1}{4}\right)$, we can write its equation as

$$y-\frac{1}{4}=k\left(x-\frac{1}{2}\right),$$

where k is the slope of the curve, i.e. the slope of the tangent. In view of the geometric meaning of derivative, the slope of the tangent is the derivative of the function $y=x^2$ at $x=\frac{1}{2}$. On using the general rule to differentiate $y=x^2$, we find $y'=2x$ for the derived function. Hence $k=y'_{x=\frac{1}{2}}=2.\frac{1}{2}=1$. The tangent to the parabola at $M_0\left(\frac{1}{2}, \frac{1}{4}\right)$ is thus given by

$$y-\frac{1}{4}=x-\frac{1}{2}$$

or

$$4x-4y-1=0\ .$$

Since the tangent to the parabola $y=x^2$ at the point $\left(\frac{1}{2}, \frac{1}{4}\right)$ has a slope of 1, it forms an angle of 45° with the Ox axis (fig. 48).

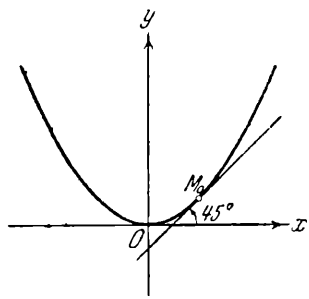


FIG. 48

(*b*) Since the tangent here is a straight line through the origin, we can write its equation as

$$y=kx$$

where k will be the derivative of the function $y=x^2$ at the point $x=0$. We find for the derivative of $y=x^2$:

$$y' = 2x\,.$$

On substituting $x=0$, we get for the slope of the tangent

$$k=0\,.$$

The equation of the tangent is therefore

$$y=0\,.$$

It will be seen that the tangent to the parabola $y=x^2$ at its vertex is the Ox axis (cf. the beginning of the present section).

3. We found in § 42 (example 5) that the derivative of the linear function $y=kx+b$ is equal to k for any x, i.e. $y'=k$. This result is obvious if we interpret the derivative geometrically as the slope of the tangent to the graph of the function. For the graph of $y=kx+b$ is a straight line of slope k, and the tangent at any point of a straight line coincides with the straight line itself; the slope of the tangent is thus a constant in this case, equal to the slope k of the straight line.

§ 44. The connexion between the existence of the derivative and the continuity of a function. Our treatment of derivatives has so far left out of account the problem of under what conditions a derivative exists at a given point. In fact, a function can only have a derivative at a given value of the argument when it is continuous at that value. This follows from the theorem below.

THEOREM. *If the function $y=f(x)$ has a derivative $f'(x_0)$ at a given point $x=x_0$, the function $f(x)$ is continuous at x_0.*

Proof. Since $f'(x_0)=\lim\limits_{\Delta x\to 0}\dfrac{\Delta y}{\Delta x}$, it follows from the definition of limit that

$$\frac{\Delta y}{\Delta x}-f'(x_0)=\alpha$$

is an infinitesimal as $\Delta x\to 0$. We obtain on solving this equation for the increment Δy

$$\Delta y=f'(x_0)\,\Delta x+\alpha\cdot\Delta x, \tag{15}$$

where $\alpha\to 0$ as $\Delta x\to 0$. The derivative $f'(x_0)$ at $x=x_0$ is a constant.

Hence $f'(x_0)$. $\Delta x \to 0$ as $\Delta x \to 0$. Thus both terms on the right-hand side of (15) tend to zero as $\Delta x \to 0$. Thus $\Delta y \to 0$ as $\Delta x \to 0$. The function is therefore continuous at $x = x_0$ (cf. the definition of continuity in § 38).

The converse is not true, because a function which is continuous at a given point does not necessarily have a derivative at that point. We shall return to this point in § 62. We shall be mainly concerned in practice, however, with continuous functions which do have a derivative at every point.

EXERCISES

On § 38.

Find:

1. $\lim\limits_{x \to 4} \dfrac{\sqrt{x} - 2x + 3}{\sqrt{25 - x^2} + 1}$. Ans. $-\dfrac{3}{4}$.

2. $\lim\limits_{x \to \frac{\pi}{4}} \dfrac{\sin 2x}{2 \cos (\pi - x)}$. Ans. $-\dfrac{1}{\sqrt{2}}$

3. $\lim\limits_{x \to 1} \dfrac{x^2 - 1}{\sqrt{x} - 1}$. Ans. 4.

4. $\lim\limits_{x \to 1} \dfrac{x^2 - \sqrt{x}}{\sqrt{x} - 1}$. Ans. 3.

On §§ 40 *and* 41.

5. Find the velocity of a particle at the end of the third second, when the distance s in metres traversed in t seconds is given by $s = 2t^3 - 3$.

Ans. 54 m/sec.

6. If a particle moves according to the law $s = t^2 - 4t + 5$, when does its velocity become zero?

Ans. At $t = 2$.

7. Find the rate of change of the function

$$y=3x^2-4x+2$$

at $x=\frac{2}{3}$.

Ans. 0.

8. Find the average rate of change of the function $y=x^2-3$ when x varies from $x=2$ to $x=3{\cdot}5$, and find the rate of change of the function at $x=2$ and $x=3{\cdot}5$.

Ans. 5·5; 4; 7.

On §§ 42 *and* 43.

9. $f(x)=3x^2$. Find $f'(2)$; $f'(-3)$; $f'\left(-\frac{1}{2}\right)$.

Ans. $f'(2)=12$; $f'(-3)=-18$; $f'\left(-\frac{1}{2}\right)=-3$.

10. $f(x)=\frac{x^3}{3}$. Find $f'(0)$; $f'(1)$; $f'(-2)$; $f'(\sqrt{2})$.

Ans. $f'(0)=0$; $f'(1)=1$; $f'(-2)=4$; $f'(\sqrt{2})=2$.

11. $f(x)=\frac{1}{x}$; find $f'(2)$; $f'\left(-\frac{1}{2}\right)$; $f'(\sqrt{2})$.

Ans. $f'(2)=-\frac{1}{4}$; $f'\left(-\frac{1}{2}\right)=-4$; $f'(\sqrt{2})=-\frac{1}{2}$.

12. Find the derived function of $y=x^2+1$.

Ans. $y'=2x$.

13. Find the derived function of

$$y=2x^2-4x+2.$$

Ans. $y'=4x-4$.

14. Find the derivative of the function $y=\frac{1}{x^2}$.

Ans. $y'=-\frac{2}{x^3}$.

By applying the differentiation rule, find the slopes of the following curves at the points indicated and draw the tangents at these points:

15. $y=x^2-4$ at the point whose abscissa is 2.

Ans. 4.

16. $y=6-x^2$ at the point whose abscissa is 1.

Ans. -2.

17. $y=\dfrac{2}{x}$ at the point whose abscissa is -2.

Ans. $-\dfrac{1}{2}$.

18. $y=x-x^2$ at the point whose abscissa is 0.

Ans. 1.

19. $y=\dfrac{1}{x-1}$ at the point whose abscissa is 3.

Ans. $-\dfrac{1}{4}$.

20. $y=\dfrac{1}{2}x^2$ at the point whose abscissa is 4.

Ans. 4.

21. $y=x^2-2x+3$ at the point whose abscissa is 1.

Ans. 0.

22. $y=9-x^2$ at the point whose abscissa is -3.

Ans. 6.

23. Find the point on the curve $y=3x^3-4x^2$ at which the tangent forms an angle of $\dfrac{\pi}{4}$ radians with Ox.

Ans. $(1, -1)$; $\left(-\dfrac{1}{9}, -\dfrac{13}{243}\right)$.

24. If a tangent to the curve $y=-2x^2+8x-9$ is parallel to Ox, find the co-ordinates of the point of contact.

Ans. $(2, -1)$.

25. Find the equation of the tangent to the curve $y=4x^2+4x-3$ at the point whose abscissa is -1.

Ans. $4x+y+7=0$.

26. Find the equation of the tangent to the hyperbola $y=\dfrac{1}{x}$ at the point whose abscissa is 1.

Ans. $x+y-2=0$.

27. At what point is the tangent to the parabola $y=x^2-4x+3$ *a*) parallel to Ox; (*b*) at an angle of 45° to Ox?

Ans. (*a*) $(2, -1)$; (*b*) $\left(\dfrac{5}{2}, -\dfrac{3}{4}\right)$.

28. Show that there is no point on the hyperbola $y=\dfrac{1}{x}$ at which the tangent is parallel to Ox.

29. At what point of the parabola $y=x^2$ is the tangent: (*a*) parallel to the straight line $4x-y+1=0$; (*b*) perpendicular to this line?

Ans. (*a*) $(2, 4)$; (*b*) $\left(-\dfrac{1}{8}, \dfrac{1}{64}\right)$.

30. What are the angles between the tangents to the parabolas $y=\sqrt{2x}$, $y=\dfrac{1}{2}x^2$ at their points of intersection?

Ans. 90°; arc tan $\dfrac{3}{4}$.

31. What is the angle between the tangents to the parabola $y=\sqrt{x}$ and the hyperbola $y=\dfrac{1}{x}$ at their point of intersection?

Ans. arc tan 3.

CHAPTER 6

BASIC FORMULAE AND RULES OF DIFFERENTIAL CALCULUS. DERIVATIVES OF ELEMENTARY FUNCTIONS

§ 45. Table of basic formulae. The differentiation rule that we gave in § 42 is fundamental inasmuch as it was deduced from the actual definition of derivative. Whilst we can use the basic rule without any particular difficulty in the case of the simpler types of function it becomes extremely unwieldy when we want to differentiate a complicated function or an expression consisting of a combination such as a sum, product or quotient of functions. Hence our obvious course is to start from the general rule and establish once and for all time special subsidiary rules for differentiating a sum, product and quotient of functions, together with a differentiation rule for so-called functions of a function (see § 52).

We first give the table of rules and basic formulae for differentiation. The symbol ( )' indicates that the derivative is taken of the expression inside the brackets.

Table of rules and basic formulae for differentiation

If u and v are functions of x whilst c is constant, we have:

I. $(u \pm v)' = u' \pm v'$.

II. $(u \cdot v)' = u'v + v'u$.

III. $(c \cdot u)' = cu'$.

IV. $\left(\frac{u}{c}\right)' = \frac{u'}{c}$.

V. $\left(\frac{u}{v}\right)' = \frac{u'v - v'u}{v^2}$ $(v \neq 0)$.

VI. *If* $y=f(u)$, where $u=\varphi(x)$, then

$$y'_x=f'(u)\,\varphi'(x)=y'_u\cdot u'_x.$$

VII. $(c)'=0.$

VIII. $(x)'=1.$

IX. $(x^a)'=ax^{a-1}$ (here and below, a is constant).

X. $(\sin x)'=\cos x.$

XI. $(\cos x)'=-\sin x.$

XII. $(\tan x)'=\dfrac{1}{\cos^2 x}=\sec^2 x.$

XIII. $(\cot x)'=-\dfrac{1}{\sin^2 x}=-\operatorname{cosec}^2 x.$

XIV. $(\log_a x)'=\dfrac{1}{x}\log_a e$ *.

XV. $(\log_e x)'=(\ln x)'=\dfrac{1}{x}.$

XVI. $(a^x)'=a^x\ln a.$

XVII. $(e^x)'=e^x.$

XVIII. $(\arcsin x)'=\dfrac{1}{\sqrt{1-x^2}}.$

XIX. $(\arccos x)'=-\dfrac{1}{\sqrt{1-x^2}}.$

XX. $(\arctan x)'=\dfrac{1}{1+x^2}.$

XXI. $(\operatorname{arc\,cot} x)'=-\dfrac{1}{1+x^2}.$

---

* See below, § 55, regarding the number e.

The aim of the present chapter is to deduce the rules and formulae given in the above table and to give the reader a grasp of the technique of differentiation, i.e. the technique of applying the basic formulae for differentiating complicated expressions.

§ 46. **Derivative of a constant.** We have agreed to regard a constant c as a variable which takes a single unique value c throughout its variation (§ 29, no. 2).

We know from analytic geometry that the equation $y=c$ represents a straight line parallel to the x-axis (or coinciding with Ox if $c=0$). Whatever the abscissa x of a point of this straight line, the corresponding ordinate y is equal to the number c (fig. 49). Hence, if we regard the constant c as a variable, we can rep-

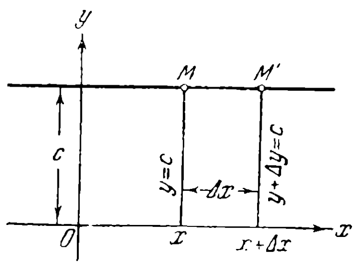


FIG. 49

resent it as the function $y=c$, taking the same value c for any argument x.

We take an arbitrary value of x and give it the increment Δx. Since the function y retains the same value c for any value of x, we have $y+\Delta y=c$ (fig. 49). Hence $\Delta y=c-y$, i.e.

$$\Delta y=c-c=0$$

and

$$\frac{\Delta y}{\Delta x}=\frac{0}{\Delta x}=0.$$

Consequently,

$$\lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x}=\lim_{\Delta x\to 0}0=0,$$

i.e.

$$(c)'=0. \tag{VII}$$

Thus *the derivative of a constant is zero.*

This result may easily be explained from the geometrical point of view. Geometrically speaking, the derivative is the slope of the tangent to a curve at a given point. But the curve in the present case is a straight line, so that the tangent coincides with the curve itself. In addition, $y=c$ is parallel to the x-axis and forms an angle of $0°$ with it. Hence the tangent is also parallel to Ox and its slope is zero:

$$y'=(c)'=\tan 0°=0 .$$

§ 47. Derivative of the function $y=x$. Following the general rule for differentiation (§ 42), we obtain:

1) $y=x$.

2) $y+\Delta y=x+\Delta x$.

3) $\Delta y=\Delta x$.

4) $\dfrac{\Delta y}{\Delta x}=\dfrac{\Delta x}{\Delta x}=1$.

5) $\lim\limits_{\Delta x\to 0}\dfrac{\Delta y}{\Delta x}=\lim\limits_{\Delta x\to 0} 1=1$.

Thus

$$(x)'=1 . \tag{VIII}$$

Our result is in accordance with the geometrical meaning of the derivative, for the function $y=x$ is represented geometrically by the bisector of the first quadrant, and since the tangent to a straight line at any point coincides with the line itself, the slope of the tangent to $y=x$ at any point is 1.

§ 48. Derivative of the product of two or more functions. 1. Suppose $y=uv$, where u and v are functions of the argument x, and let u and v have derivatives u' and v' for the value of x in question. We shall find the derivative of the function $y=uv$, which is the product of functions u and v.

We follow the general rule for differentiation and give the value of x concerned the increment Δx. The functions u, v and y receive the increments Δu, Δv and Δy respectively and pass from their initial values u, v and y to the values $u+\Delta u$, $v+\Delta v$ and $y+\Delta y$. The new values of the functions are connected by the relationship

$$y+\Delta y=(u+\Delta u)\,(v+\Delta v)\,,$$

or

$$y+\Delta y=uv+\Delta u\cdot v+\Delta v\cdot u+\Delta u\cdot\Delta v\,.$$

We subtract from this the equation $y=uv$ and find for Δy

$$\Delta y=\Delta u\cdot v+\Delta v\cdot u+\Delta u\cdot\Delta v\,.$$

We form the ratio $\dfrac{\Delta y}{\Delta x}$:

$$\frac{\Delta y}{\Delta x}=\frac{\Delta u}{\Delta x}\cdot v+\frac{\Delta v}{\Delta x}\cdot u+\frac{\Delta u}{\Delta x}\cdot\Delta v\,. \tag{1}$$

When $\Delta x\to 0$ the original value of x at which we are finding the derivative remains unchanged. With x unchanged and $\Delta x\to 0$, the functions u and v remain constant. Hence

$$\lim_{\Delta x\to 0} u=u \quad\text{and}\quad \lim_{\Delta x\to 0} v=v\,.$$

Moreover, since by hypothesis functions u and v have derivatives at the value of x concerned, they are continuous for this x (§ 44). Their increments Δu and Δv thus tend to zero as $\Delta x\to 0$ (§ 38). Consequently

$$\lim_{\Delta x\to 0}\left(\frac{\Delta u}{\Delta x}\cdot\Delta v\right)=\lim_{\Delta x\to 0}\frac{\Delta u}{\Delta x}\lim\Delta v=u'\cdot 0=0\,.$$

On now applying the limit theorems for an algebraic sum of functions and a product, we obtain

$$\lim_{\Delta x\to 0}\left(\frac{\Delta u}{\Delta x}\cdot v+\frac{\Delta v}{\Delta x}\,u+\frac{\Delta u}{\Delta x}\,\Delta v\right)=u'v+v'u\,.$$

In view of equation (1), we now conclude that

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = u'v + v'u ,$$

i.e.

$$y' = (uv)' = u'v + v'u . \tag{II}$$

Thus *the derivative of the product of two functions is equal to the product of the derivative of the first function with the second plus the product of the derivative of the second function with the first* (assuming that each function has a derivative).

2. Now suppose that $y = u_1 u_2 u_3 u_4$, where u_1, u_2, u_3 and u_4 are functions of the argument x having derivatives u'_1, u'_2, u'_3 and u'_4 at the value of x concerned. We can write our given product of four functions as the product of two factors: $y = (u_1 u_2)(u_3 u_4)$. We find on using rule (II) above

$$\begin{aligned} y' &= (u_1 u_2)'(u_3 u_4) + (u_1 u_2)(u_3 u_4)' \\ &= (u'_1 u_2 + u'_2 u_1)(u_3 u_4) + (u_1 u_2)(u'_3 u_4 + u'_4 u_3) \\ &= u'_1 (u_2 u_3 u_4) + u'_2 (u_1 u_3 u_4) + u'_3 (u_1 u_2 u_4) + u'_4 (u_1 u_2 u_3). \end{aligned}$$

It may easily be seen, on generalizing the rule for the product of n functions, that

$$\begin{aligned} (u_1 u_2 u_3 \ldots u_n)' &= u'_1 (u_2 u_3 \ldots u_n) + u'_2 (u_1 u_3 \ldots u_n) + \\ &+ u'_3 (u_1 u_2 u_4 \ldots u_n) + \ldots + u'_n (u_1 u_2 \ldots u_{n-1}). \end{aligned} \tag{II*}$$

3. If $y = cu$, where the factor c is constant, we find on applying the differentiation rule (II) for a product

$$y' = c'u + cu' :$$

and since $(c)' = 0$ (formula VII), we finally obtain

$$(cu)' = cu', \tag{III}$$

i.e. *the derivative of the product of a constant and a function is equal to the product of the constant and the derivative of the function.*

This rule is often expressed as follows: *a constant factor can be taken outside the sign of differentiation.*

§ 49. **Derivative of a positive integral power.** We write the power function $y=x^n$, where n is a positive integer, as the product

$$y=\underbrace{x \cdot x \ldots x}_{n}.$$

We now make use of rule (II*) and formula (VIII) to obtain

$$y'=1 \cdot \underbrace{(x \cdot x \ldots x)}_{n-1}+1 \cdot \underbrace{(x \cdot x \ldots x)}_{n-1}+\ldots+1 \cdot \underbrace{(x \cdot x \ldots x)}_{n-1},$$

or

$$y'=\underbrace{x^{n-1}+x^{n-1}+\ldots+x^{n-1}}_{n}=nx^{n-1}.$$

Thus *the derivative of the positive integral power* x^n *is equal to the index of the power multiplied by the base* x *raised to the original power minus one,* i.e.

$$(x^n)'=nx^{n-1}. \tag{IX*}$$

We have only proved formula (IX) of the table of basic formulae (§ 45) for a *positive integral* power n. We show in § 57 that it also remains true for any value of n (fractional, negative, rational, irrational). To prevent delay, however, in obtaining some practice in differentiation, we shall make immediate use of formula (IX*) for any index n.

Example 1. Differentiate the function $y=x^5$.

Solution. Here $n=5$. Hence we obtain from formula (IX*) $y'=5x^4$.

Example 2. Differentiate the function $y=\sqrt{x}$.

Solution. Since $y=\sqrt{x}=x^{\frac{1}{2}}$, we obtain from formula (IX) $y'=\frac{1}{2}x^{-\frac{1}{2}}$, or

$$(\sqrt{x})'=\frac{1}{2\sqrt{x}}. \tag{IX**}$$

Cases where we find ourselves needing to differentiate $\sqrt{x}$ are very common in practice. It is therefore advisable to memorize

formula (IX**) rather than keep returning to the general formula (IX).

Example 3. Differentiate the function $y=\frac{3}{\sqrt[3]{x^2}}$.

Solution. On writing the function in the form

$$y=3\cdot x^{-\frac{2}{3}},$$

we find by rule (III) and formula (IX)

$$y'=3\cdot\left(-\frac{2}{3}\right)x^{-\frac{5}{3}}=-\frac{2}{x\sqrt[3]{x^2}}.$$

§ 50. Derivative of an algebraic sum of functions. Let the given function be

$$y=u\pm v,$$

where u and v are functions of the argument x which have derivatives u' and v' at the value of x in question.

Following the general rule for differentiation, we get

1) $$y=uy\pm v.$$

2) We give the argument x the increment Δx. Then v, u and y receive respectively the increments Δu, Δv and Δy. The new values of these functions: $u+\Delta u$, $v+\Delta v$, $y+\Delta y$ are connected by the relationship

$$y+\Delta y=(u+\Delta v)\pm(v+\Delta v)$$

(3)
$$\begin{array}{r} -\,y+\Delta y=u+\Delta u\pm v\pm\Delta v\\ \underline{\;\;y\qquad\;=u\qquad\;\pm v\qquad\;}\\ \Delta y=\Delta u\pm\Delta v\;. \end{array}$$

(4) $$\frac{\Delta y}{\Delta x}=\frac{\Delta u}{\Delta x}\pm\frac{\Delta v}{\Delta x}.$$

5) Since by hypothesis $\lim\limits_{\Delta x\to}\frac{\Delta u}{\Delta x}=u'$, $\lim\limits_{\Delta x\to 0}\frac{\Delta v}{\Delta x}=v'$, we obtain from the limit theorem for a sum

$$y'=u'\pm v'.$$

This result can be extended without difficulty to the sum of any number of terms.

Example. Find the derivative of the function $y=x^2+\frac{1}{x}$.

Solution. We have by rule (I)

$$y'=(x^2)'+\left(\frac{1}{x}\right)',$$

and since by formula (IX)

$$(x^2)'=2x, \quad \left(\frac{1}{x}\right)'=(x^{-1})'=-1\cdot x^{-2}=-\frac{1}{x^2},$$

we find

$$y'=2x-\frac{1}{x^2}.$$

§ 51. Derivative of a quotient. Let us be given the function

$$y=\frac{u}{v},$$

where u and v are functions of the argument x which have derivatives u' and v' at the given value of x. We assume, furthermore, that v differs from zero for the value of x in question.

We obtain on giving the argument x the increment Δx:

$$y+\Delta y=\frac{u+\Delta u}{v+\Delta v}.$$

We subtract from this equation the equation $y=\frac{u}{v}$ and find Δy as

$$\Delta y=\frac{u+\Delta u}{v+\Delta v}-\frac{u}{v}=\frac{\Delta u\cdot v-\Delta v\cdot u}{v^2+\Delta v\cdot v}.$$

Further,

$$\frac{\Delta y}{\Delta x}=\frac{\frac{\Delta u}{\Delta x}\cdot v-\frac{\Delta v}{\Delta x}\cdot u}{v^2+\Delta v\cdot v}.$$

By hypothesis, $\lim\limits_{\Delta x\to 0}\frac{\Delta u}{\Delta x}=u'$, $\lim\limits_{\Delta x\to 0}\frac{\Delta v}{\Delta x}=v'$. Also, u and v are constant for the given x and $\Delta x\to 0$; $v^2\neq 0$ and $\Delta v\to 0$ when $\Delta x\to 0$. Consequently, we have by the limit theorem

$$y'=\lim_{\Delta x\to 0}\frac{\Delta y}{\Delta x}=\frac{\lim\limits_{\Delta x\to 0}\frac{\Delta u}{\Delta x}\cdot v-\lim\limits_{\Delta x\to 0}\frac{\Delta v}{\Delta x}\cdot u}{v^2+0\cdot v},$$

or

$$\left(\frac{u}{v}\right)'=\frac{vu'-uv'}{v^2}. \tag{V}$$

Relationship (V) may be stated as follows: *to find the derivative of a quotient, multiply the denominator by the derivative of the numerator, subtract from this the product of the numerator and the derivative of the denominator, then divide the result by the square of the denominator.*

In particular, if $u=c$, where c is constant, we have by formula (V)

$$y'=\frac{c'v-v'c}{v^2},$$

or, since $c'=0$,

$$y'=-\frac{cv'}{v^2}. \tag{V*}$$

On the other hand, if $v=c$, then $y=\frac{u}{c}$. A fraction whose denominator is constant should never be differentiated by using formula (V) as we only need to notice that $\frac{u}{c}=\frac{1}{c}\cdot u$ and it follows at once from (III) that

$$\left(\frac{1}{c}u\right)'=\frac{1}{c}u'=\frac{u'}{c}. \tag{VI}$$

Example 1. Find the derivative of the function $y=\frac{1-x^3}{1+x^3}$.

Solution. We have by rule (V)

$$y' = \frac{(1-x^3)'(1+x^3)-(1+x^3)'(1-x^3)}{(1+x^3)^2}.$$

Hence, on applying rule (I) and formula (IX), we obtain

$$y' = \frac{-3x^2(1+x^3)-3x^2(1-x^3)}{(1+x^3)^2},$$

or, after simplifying,

$$y' = -\frac{6x^2}{(1+x^3)^2}.$$

Example 2. Find the derivative of $y=\frac{2}{x^4}$.

Solution. We find by formula (V*)

$$y' = -\frac{8x^3}{x^8} = -\frac{8}{x^5}.$$

Example 3. Find the derivative of $y=\frac{3x^3+5x}{7}$.

Solution. We obtain from formulae (IV), (III) and (IX)

$$y' = \frac{9x^2+5}{7}.$$

§ **52. Functions of a function and their derivatives.** Suppose we are given the function $y=\log\sin x$. Here we have a logarithmic function whose argument is not the independent variable x but the function $\sin x$ of this variable. A function of this type is called a *function of a function* (or *composite function*).

Let u denote $\sin x$. Then $y=\log u$, where $u=\sin x$. We have introduced an auxiliary variable u and written the relationship between y and x as a relationship between y and u, and one between u and x.

Similarly, if we write u for $1+x^2$ in the relationship $y=\sqrt{(1+x^2)}$, we can represent this latter as a relationship between y and u, and one between u and x.

In general, if

$$y = f(u), \tag{a}$$

where

$$u = \varphi(x), \tag{b}$$

we say that relationships (a) and (b) define y as a composite function of x or as a function of the function u.

If we replace u by $\varphi(x)$ in the relationship $y = f(u)$, we obtain

$$y = f[\varphi(x)].$$

This way of writing y makes it clear that the argument of the function f is not the independent variable x but the function $\varphi(x)$.

We again take the function

$$y = \sqrt{1+x^2} = (1+x^2)^{\frac{1}{2}}.$$

What rule can be used for finding the derivative of this function? We cannot use our rule for differentiating a power series since this was obtained for a function of the form x^n, i.e. a power of the argument x itself, whereas we are concerned in $y = (1+x^2)^{\frac{1}{2}}$ with a power of the function $1+x^2$. Since formula (IX) is obviously inapplicable in a case like this, we must either return directly to the general rule for finding a derivative (§ 42), or we must find a new rule for differentiating a function of a function. The aim of the present section is to deduce the special rule for differentiation of a function of a function.

Consider the function

$$y = f(u), \text{ where } u = \varphi(x).$$

We shall assume that the function $u = \varphi(x)$ has a derivative for the given value of x, whilst $f(u)$ has a derivative $f'(u)$ for the value of u corresponding to the given x. We say in this case that u has a derivative with respect to the variable x and use the notation u'_x. Similarly, $f'(u)$ is the derivative of y with respect to the variable u, i.e. we write $f'(u) = y'_u$.

We shall deduce the required formula by following the general rule for differentiation:

1) We assign the given x the increment Δx. Then the function $u=\varphi(x)$ receives the increment Δu, which in turn defines the increment Δy of the function $y=f(u)$. The result of the first operation is therefore as follows:

$$u+\Delta u=\varphi(x+\Delta x)\,; \qquad y+\Delta y=f(u+\Delta u).$$

2) We find the increments Δu and Δy as follows:

$$\begin{array}{rl} -\,u+\Delta u & =\varphi(x+\Delta x) \\ u & =\varphi(x) \\ \hline \Delta u & =\varphi(x+\Delta x)-\phi(x) \end{array} \qquad \begin{array}{rl} -\,y+\Delta y & =f(u+\Delta u) \\ y & =f(u) \\ \hline \Delta y & =f(u+\Delta u)-f(u). \end{array}$$

3) To find the derivative y'_x we have to find the limit of the ratio $\dfrac{\Delta y}{\Delta x}$ as $\Delta x\to 0$. But the increment Δy is given in terms of the increment Δu, and not directly in terms of Δx. Thus, to find the limit of the ratio

$$\frac{\Delta y}{\Delta x}=\frac{f(u+\Delta u)-f(u)}{\Delta x}$$

as $\Delta x\to 0$ we must first write this ratio in the following form:

$$\frac{\Delta y}{\Delta x}=\frac{f(u+\Delta u)-f(u)}{\Delta u}\cdot\frac{\Delta u}{\Delta x}=\frac{\Delta y}{\Delta u}\cdot\frac{\Delta u}{\Delta x}. \tag{2}$$

4) Now let Δx tend to zero.

Since the function u has by hypothesis a derivative u'_x for the given x, u must be a continuous function for this x (§ 44), so that $\Delta u\to 0$ as $\Delta x\to 0$. Also by hypothesis, the function $y=f(u)$ has a derivative $f'(u)=y'_u$ for the u corresponding to the given x. Consequently there exists $\lim\limits_{\Delta u\to 0}\dfrac{\Delta y}{\Delta u}=y'_u$. We obtain by applying the theorem concerning the limit of a product:

$$\lim_{\Delta x\to 0}\left(\frac{\Delta y}{\Delta u}\cdot\frac{\Delta u}{\Delta x}\right)=\lim_{\Delta x\to 0}\frac{\Delta y}{\Delta u}\cdot\lim_{\Delta x\to 0}\frac{\Delta u}{\Delta x}=\lim_{\Delta u\to 0}\frac{\Delta y}{\Delta x}\cdot u'_x=y'_u\cdot u'_x\,.$$

We now conclude, in view of equation (2), that

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = y'_u \cdot u'_x . \tag{2/2}$$

i.e.

$$y'_x = y'_u \cdot u'_x . \tag{VI}$$

In other words, *if $y=f(u)$, where $u=\varphi(x)$, the derivative of y with respect to the variable x is equal to the derivative of y with respect to the variable u multiplied by the derivative of u with respect to the variable x* *.

On replacing y'_u by $f'(u)$ and u'_x by $\varphi'(x)$, formula (VI) can be written

$$y'_x = f'(u) \cdot \varphi'(x) .$$

Example 1. $$y=\sqrt{a^2-x^2} .$$

Solution. We are given the composite function $y=\sqrt{u}$, $=a^2-x^2$. We apply formula (VI) and get

$$y'_x = \frac{1}{2\sqrt{u}} \cdot u'_x = \frac{1}{2\sqrt{a^2-x^2}} (a^2-x^2)'$$

$$= \frac{1}{2\sqrt{a^2-x^2}} (-2x) = -\frac{x}{\sqrt{a^2-x^2}} .$$

* *Note for* lecturers. It may be observed that we have deduced formula (VI) without taking into account the possibility of the increment Δu of the function $u=\varphi(x)$ vanishing as $\Delta x \to 0$. When the limit of $\frac{\Delta y}{\Delta x}$ is sought as $\Delta x \to 0$, the *independent variable is* Δx, so that we are justified in excluding the possibility of Δx vanishing (cf. § 31). But since Δu is the increment of $u=\varphi(x)$, as $\Delta x \to 0$ Δu can take zero values. The right-hand side of equation (2) becomes meaningless with $\Delta u=0$. Hence the case when Δu takes zero values as $\Delta x \to 0$ requires special consideration, formula (VI) still remains valid in this case. But the proof of this statement lies outside the scope of an elementaıy work foı technical schools and is therefore omitted. A rigorous proof of (VI) can be found, for instance, in Luzin's *Differential Calculus*.

The process of differentiation becomes very cumbersome when written down in the above way.

We shall use the same example to show a simpler method of writing.

Let the function $y=\sqrt{a^2-x^2}$ be given. It is at once obvious that this is a function of a function of x. We replace the a^2-x^2 under the root sign by u in our heads, differentiate y with respect to $u=a^2-x^2$ and multiply by the derivative of u with respect to x, i.e. by $(a^2-x^2)'$.
Thus

$$y'=\frac{1}{2\sqrt{a^2-x^2}}(a^2-x^2)'=-\frac{x}{\sqrt{a^2-x^2}}.$$

Here we have omitted the subscript x in the symbol y'_x for the derivative. We shall also simply write y' in future when differentiating a function y with respect to the variable x.

Finally, after the reader has acquired some practice in differentiation, the writing can be further simplified. This simplification consists, as regards our present example, in writing straight down the result of differentiating $u=a^2-x^2$ instead of putting $(a^2-x^2)'$. The working now appears as follows:

$$y'=\frac{1}{2\sqrt{a^2-x^2}}\cdot(-2x)=-\frac{x}{\sqrt{a^2-x^2}}.$$

Example 2.

$$y=\sqrt[3]{1+3x^2}$$

Solution.

$$y'=[(1+3x^2)^{\frac{1}{3}}]'=\frac{1}{3}(1+3x^2)^{-\frac{2}{3}}\cdot(6x)=\frac{2x}{\sqrt[3]{(1+3x^2)^2}}.$$

[by formulae (VI), (I), (VII), (III) and (IX)].

Example 3.

$$y=\frac{a^2+x^2}{\sqrt{a^2-x^2}}.$$

Solution.

$$y'=\frac{(a^2+x^2)'\sqrt{a^2-x^2}-(\sqrt{a^2-x^2})'(a^2+x^2)}{a^2-x^2}$$

[by formula (V)]

$$=\frac{2x\sqrt{a^2-x^2}-\dfrac{1}{2\sqrt{a^2-x^2}}(-2x)(a^2+x^2)}{a^2-x^2}$$

$$=\frac{2x(a^2-x^2)+x(a^2+x^2)}{(a^2-x^2)^{\frac{3}{2}}}=\frac{3a^2x-x^3}{(a^2-x^2)^{\frac{3}{2}}}.$$

§ 53. **Limit of the ratio** $\dfrac{\sin z}{z}$ **as** $z\to 0$. We require this limit for deducing the derivatives of the trigonometric functions. We shall establish that

$$\lim_{z\to 0}\frac{\sin z}{z}=1.$$

The theorem regarding the limit of a quotient cannot be applied to the present limit since the denominator here tends to

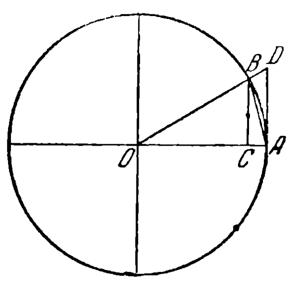


FIG. 50

zero. We therefore have to look for some other way of solving our problem. The method is as follows:

We take a circle of unit radius (fig. 50). We write z for the radian measure of the arc, i.e. the ratio of the length of arc to the

radius, so that z is numerically equal to the length of arc in the present case. We shall take z in the interval from 0 to $\frac{1}{2}\pi$ so that

$$0<z<\frac{\pi}{2}.$$

We have from fig. 50:

area $\Delta OAB <$ area sector $OAB <$ area ΔOAD or

$$\frac{1}{2}OA\cdot CB<\frac{1}{2}OA\cdot\breve{AB}<\frac{1}{2}OA\cdot AD.$$

Since $OA=1$, we have $CB=\sin z$, $\breve{AB}=z$, $AD=\tan z$. The above pair of inequalities can thus be written

$$\sin z<z<\tan z.$$

On dividing all the terms of this pair of inequalities by the positive quantity $\sin z\left(\sin z>0 \text{ since } 0<z<\frac{1}{2}\pi\right)$, we obtain

$$1<\frac{z}{\sin z}<\frac{1}{\cos z}$$

or

$$1>\frac{\sin z}{z}>\cos z.$$

Now let z tend to zero. We have $\lim\limits_{z\to 0} 1=1$ and $\lim\limits_{z\to 0}\cos z=1$. We thus have the variable $\frac{\sin z}{z}$ lying between two other variables (unity and $\cos z$) which have the same limit. It now follows from theorem 4 of § 32 that

$$\lim_{z\to 0}\frac{\sin z}{z}=1.$$

We have so far assumed that $z\to 0$ whilst remaining positive.

Now let $z \to 0$ in an arbitrary manner (i.e. taking negative as well as positive values). Suppose that z takes the negative value $-z_1$ $(z_1 > 0)$. Since $\sin z = \sin(-z_1) = -\sin z_1$, we get

$$\frac{\sin z}{z} = \frac{\sin(-z_1)}{-z_1} = \frac{-\sin z_1}{-z_1} = \frac{\sin z_1}{z_1}.$$

Hence it follows that the ratio $\frac{\sin z}{z}$ is independent of the sign of z. Its limit as $z \to 0$ must therefore be independent of the method in which z tends to zero. Thus

$$\lim_{z \to 0} \frac{\sin z}{z} = 1.$$

In order that the result obtained may be better understood, we reproduce the graphs of $y=1$, $y=\cos z$ and $y=\frac{\sin z}{z}$ (fig. 51).

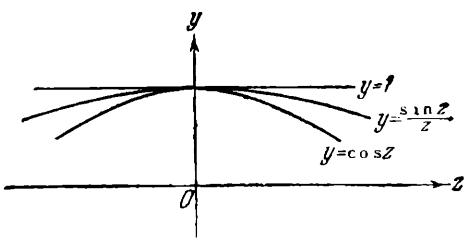


FIG. 51

In accordance with the inequalities $1 > \frac{\sin z}{z} > \cos z$, the curve representing $y = \frac{\sin z}{z}$ lies below the straight line $y=1$ and above the curve $y=\cos z$. The ordinates of these latter curves ($y=1$ and $y=\cos z$) have the limit 1. And since the curve $y=\frac{\sin z}{z}$ can neither fall below the curve $y=\cos z$ nor rise above the straight line $y=1$, it follows that

$$\lim_{z\to 0}\frac{\sin z}{z}$$

is also equal to unity.

The following table shows how the ratio $\frac{\sin z}{z}$ varies as $z\to 0$:

z	$\frac{\pi}{9}$	$\frac{\pi}{18}$	$\frac{\pi}{36}$	$\frac{\pi}{180}$	...	$\to 0$
$\frac{\sin z}{z}$	0·9798	0·9949	0.9987	0·9999	...	$\to 1$

§ 54. **Derivatives of the trigonometric functions.** 1. *Derivative of the function* $y=\sin x$. We obtain on applying the general rule for differentiation:

1) $y=\sin x.$

2) $y+\Delta y=\sin(x+\Delta x).$

3)
$$\begin{aligned} y+\Delta y&=\sin(x+\Delta x)\\ y\quad&=\sin x\\ \hline \Delta y&=\sin(x+\Delta x)-\sin x \end{aligned}$$

$$=2\cos\frac{2x+\Delta x}{2}\sin\frac{\Delta x}{2}=2\cos\left(x+\frac{\Delta x}{2}\right)\cdot\sin\frac{\Delta x}{2}.$$

4)
$$\frac{\Delta y}{\Delta x}=\frac{2\cos\left(x+\frac{\Delta x}{2}\right)\sin\frac{\Delta x}{2}}{\Delta x}=\cos\left(x+\frac{\Delta x}{2}\right)\cdot\frac{\sin\frac{\Delta x}{2}}{\frac{\Delta x}{2}}$$

$$=\cos\left(x+\frac{\Delta x}{2}\right)\cdot\frac{\sin z}{z},$$

where $z=\frac{1}{2}\Delta x$. As $\Delta x\to 0$ the variable z also tends to 0.

5) Since the function $\cos(x+\frac{1}{2}\Delta x)$ is continuous for any Δx, we have

$$\lim_{\Delta x \to 0} \cos\left(x+\frac{\Delta x}{2}\right)=\cos\left(x+\frac{0}{2}\right)=\cos x\ ; \quad \lim_{z\to 0} \frac{\sin z}{z}=1.$$

Thus

$$\lim_{\Delta x\to 0} \frac{\Delta y}{\Delta x}=y'=\lim_{\Delta x\to 0}\left[\cos\left(x+\frac{\Delta x}{2}\right)\cdot\frac{\sin z}{z}\right]=\cos x.$$

Consequently,

$$(\sin x)'=\cos x\ . \tag{X}$$

2. *Derivative of the function cos x.* We can write $\cos x$ in the form $\cos x=\sin\left(\frac{1}{2}\pi-x\right)$.

If we now differentiate $\sin\left(\frac{1}{2}\pi-x\right)$ as a function of a function of x, we obtain

$$(\cos x)'=\cos\left(\frac{\pi}{2}-x\right)\cdot\left(\frac{\pi}{2}-x\right)'=-\sin x,$$

since

$$\cos\left(\frac{1}{2}\pi-x\right)=\sin x \quad \text{and} \quad \left(\frac{1}{2}\pi-x\right)'=-1.$$

Hence

$$(\cos x)'=-\sin x. \tag{XI}$$

3. *Derivative of the function tan x.* Since

$$\tan x=\frac{\sin x}{\cos x},$$

we get from formula (V) for the derivative of a quotient:

$$(\tan x)'=\left(\frac{\sin x}{\cos x}\right)'=\frac{\cos x\cdot\cos x-(-\sin x)\cdot\sin x}{\cos^2 x}$$

$$=\frac{\cos^2 x+\sin^2 x}{\cos^2 x}=\frac{1}{\cos^2 x}.$$

Hence

$$(\tan x)' = \frac{1}{\cos^2 x}. \qquad \text{(XII)}$$

4. *Derivative of the function cot x.* Since

$$\cot x = \frac{\cos x}{\sin x},$$

we find by the formula for the derivative of a quotient

$$(\cot x)' = \left(\frac{\cos x}{\sin x}\right)' = \frac{(-\sin x)\cdot \sin x - \cos x \cos x}{\sin^2 x}$$

$$= -\frac{\sin^2 x + \cos^2 x}{\sin^2 x} = -\frac{1}{\sin^2 x}.$$

Hence

$$(\cot x)' = -\frac{1}{\sin^2 x}. \qquad \text{(XIII)}$$

Examples. Differentiate the following functions:

1. $y = \sin 3x^2$.

Solution.

$$\left.\begin{aligned} y' &= \cos 3x^2 \cdot (3x^2)' \\ &= 6x \cdot \cos 3x^2. \end{aligned}\right\} \text{(by formulae X and VI)}$$

2. $y = \tan \sqrt{1 - x^3}$.

Solution.

$$\left.\begin{aligned} y' &= \frac{1}{\cos^2 \sqrt{1-x^3}} \cdot \sqrt{(1-x^3)}' \\ &= \frac{1}{\cos^2 \sqrt{1-x^3}} \cdot \frac{1}{2\sqrt{1-x^3}} (1-x^3)' \\ &= -\frac{3x^2}{2\sqrt{1-x^3}\cdot \cos^2 \sqrt{1-x^3}} \end{aligned}\right\} \text{(by formulae XII and VI)}$$

3. $y=\cos^3 x$.

Solution.

$$\left.\begin{aligned} y' &= 3(\cos x)^2 \cdot (\cos x)' \\ &= -3\sin x \cdot \cos^2 x \end{aligned}\right\} \text{(by formulae IX and VI).}$$

§ 55. Conversion modulus from one system of logarithms to another. The number e. Natural logarithms. Conversion from natural to common logarithms and vice versa. 1. Suppose we know that p is the logarithm of a number N to the base a ($a>0$, $a \neq 1$), i.e.

$$\log_a N = p\,,$$

then

$$a^p = N\,.$$

We take the logarithm to base b of both sides of this equation ($b>0$, $b \neq 1$), obtaining

$$p \log_b a = \log_b N\,.$$

We see from this that, to find the logarithm to base b of a number N whose logarithm to base a is already known, we multiply the known logarithm p by $\log_b a$. The factor $\log_b a$ is known as the *conversion modulus* from the system of logarithms to base a to the system to base b.

We show that $\log_a b = \dfrac{1}{\log_b a}$. For let $\log_a b = l$, then $al=b$ and $l \log_b a = \log_b b = 1$, whence we have

$$\log_a b = l = \frac{1}{\log_b a}\,.$$

2. Differentiation of logarithmic functions is based on the preliminary task of finding the limit of the function $(1+\alpha)^{\frac{1}{\alpha}}$ as $\alpha \to 0$. Comprehensive courses of mathematical analysis prove the existence of this limit. We confine ourselves in the present elementary course to some brief remarks suggesting the existence of this limit but lacking the strength of an actual proof.

We observe how the function

$$z=(1+\alpha)^{\frac{1}{\alpha}}$$

varies as α tends to zero. This is done by drawing up the following table giving the values of z corresponding to diminishing values of the argument α:

α	z	α	z
10	1·0096		
5	1·4310		
2	1·7320		
1	2·0000		
0·5	2·2500	−0·5	4·0000
0·1	2·5937	−0·1	2·8680
0·01	2·7048	−0·01	2·7320
0·001	2·7169	−0·001	2·7195

It is clear from this table that the value of $(1+\alpha)^{1/\alpha}$ approaches a number 2·71... as the absolute value of α decreases. A more exact evaluation shows that

$$\lim_{\alpha\to 0}(1+\alpha)^{\frac{1}{\alpha}}=2.718281828459045\ldots$$

The limit is seen to be an irrational number, which is generally denoted by the letter e.

Thus

$$\lim_{\alpha\to 0}(1+\alpha)^{\frac{1}{\alpha}}=e.$$

The number e is of great importance in analysis. Whereas 10 proves to be a convenient base for logarithms used in practical work, a system of logarithms with the base e is more suitable for theoretical work.

In future we shall denote the natural logarithm of a number N, i.e. $\log_e N$, by the symbol $\ln N$.

3. Let $a=e$ and $b=10$. Then the formula obtained in 1. above gives us

$$\log_{10} N=\log_e N\cdot\log_{10} e=\log_e N\cdot\frac{1}{\log_e 10}.$$

Logarithms to the base e are known as *natural* or *Neperian* logarithms. The conversion modulus from natural to common logarithms is equal to 0·4329..., and

$$\log_{10} e=\frac{1}{\log_e 10}\approx 0{\cdot}43429.$$

Hence

$$\log_{10} N \approx \log_e N\cdot 0.43429.$$

Similarly

$$\log_e N=\log_{10} N\cdot\log_e 10=\log_{10} N\cdot\frac{1}{\log_{10} e}\approx$$

$$\approx\frac{\log_{10} N}{0\cdot 43429}\approx\log_{10} N\cdot 2\cdot 3021$$

§ **56. Derivatives of logarithmic functions.** *The derivative of the function $y=\log_a x$ is equal to the reciprocal of the argument x multiplied by the modulus of the system of logarithms, i.e.*

$$(\log_a x)'=\frac{1}{x}\log_a e. \qquad \text{(XIV)}$$

For the case $a=e$ we have $\log_a e=\log_e e=1$, and formula becomes

$$(\ln x)'=\frac{1}{x}, \qquad \text{(XV)}$$

i.e. *the derivative of the function $y=\ln x$ is equal to the reciprocal of the argument x.*

Formula (XV) is simpler than (XIV). This remark reveals one of the valuable properties of natural logarithms from the point of view of mathematical analysis.

The proof of formula (XIV) is rather difficult and really lies outside the scope of the present course. Nevertheless, it is given below for the more advanced readers.

Following the general rule for finding a derivative, we have

1) $y=\log_a x$.

2) $y+\Delta y=\log_a (x+\Delta x)$.

3) $y+\Delta y=\log_a (x+\Delta x)$

$$\begin{array}{rl} - & \\ y & =\log_a x \\ \hline \end{array}$$

$$\Delta y=\log_a (x+\Delta x)-\log_a x=\log_a \frac{x+\Delta x}{x}$$

$$=\log_a\left(1+\frac{\Delta x}{x}\right).$$

4) $$\frac{\Delta y}{\Delta x}=\frac{\log_a\left(1+\frac{\Delta x}{x}\right)}{\Delta x}=\frac{1}{\Delta x}\log_a\left(1+\frac{\Delta x}{x}\right)=\log_a\left(1+\frac{\Delta x}{x}\right)^{\frac{1}{\Delta x}}.$$

To find the limit of $\frac{\Delta y}{\Delta x}$ as $\Delta x\to 0$, we transform the expression obtained as follows:

$$\frac{\Delta y}{\Delta x}=\log_a\left[\left(1+\frac{\Delta x}{x}\right)^{\frac{x}{\Delta x}}\right]^{\frac{1}{x}}=\frac{1}{x}\log_a\left(1+\frac{\Delta x}{x}\right)^{\frac{x}{\Delta x}}$$

or

$$\frac{\Delta y}{\Delta x}=\frac{1}{x}\log_a (1+\alpha)^{\frac{1}{\alpha}},$$

where we have set $\alpha=\frac{\Delta x}{x}$.

5) When $\Delta x\to 0$, $\frac{1}{x}$ remains constant and so $\alpha=\frac{\Delta x}{x}\to 0$ (since $\Delta x\to 0$ whilst x remains constant). If we write $(1+\alpha)^{1/\alpha}=u$ then

$$\log_a (1+\alpha)^{1/\alpha}=\log_a u\,,\quad \text{where}\quad u=(1+\alpha)^{1/\alpha}.$$

When $\Delta x\to 0$, we have

$$\lim_{\Delta x\to 0} u=\lim_{\Delta x\to 0}(1+\alpha)^{1/\alpha}=\lim_{\alpha\to 0}(1+\alpha)^{1/\alpha}=e\,.$$

The function $\log_a u$ is continuous for any value of the argument u and therefore, in particular, for $u=e$. Hence

$$\lim_{\Delta x \to 0} \log_a u = \lim_{u \to e} \log_a u = \log_a e.$$

Consequently

$$\lim_{\Delta x \to 0} \left[\frac{1}{x} \log_a (1+\alpha)^{\frac{1}{\alpha}} \right] = \frac{1}{x} \log_a e,$$

i.e.

$$\lim_{\Delta x \to 0} \frac{\Delta y}{\Delta x} = y' = \frac{1}{x} \log_a e.$$

Examples. Differentiate the following functions:

1. $y = \ln (x^3+2)$.

Solution.

$$\left. y' = \frac{1}{x^3+2} \cdot (x^3+2)' = \frac{3x^2}{x^3+2} \right\} \text{(by formulae XV and VI).}$$

2. $y = \ln \sqrt{1-x^2}$.

Solution.

$$y' = \frac{1}{\sqrt{1-x^2}} \cdot (\sqrt{1-x^2})'$$

$$= \frac{1}{\sqrt{1-x^2}} \cdot \frac{1}{2\sqrt{1-x^2}} (-2x) = \frac{x}{x^2-1}.$$

This example can be solved in another way, viz.

$$y = \ln \sqrt{1-x^2} = \frac{1}{2} \ln (1-x^2),$$

whence

$$\left. \begin{aligned} y' &= \frac{1}{2} [\ln (1-x^2)]' \\ &= \frac{1}{2} \cdot \frac{1}{1-x^2} (-2x) = \frac{x}{x^2-1} \end{aligned} \right| \text{(by formulae III, VI, XV).}$$

§ 57. Derivative of a power to any index. When deducing formulae for the derivatives of different functions in previous articles we have simultaneously proved that the derivative of the function in question exists. There are still several functions whose derivatives we need to find. But simultaneous proof of the existence of the derivative is rather difficult in these last cases, and we must confine ourselves in the present course merely to finding the formulae for differentiation whilst taking the existence of the derivatives for granted.

Let

$$j=x^a,$$

where a is any real number and $x>0$. As we have just mentioned it will be assumed that the function y has a derivative y'.

Taking natural logarithms of both sides of the equation $y=x^a$, we obtain

$$\ln y=a\ln x.$$

The function $\ln y$ is a function of a function of x, since the natural logarithm is taken of y, whilst y is a function of x. Hence we obtain from formulae (XV) and (VI) $(\ln y)'=\frac{1}{y}\cdot y'$.

The derivative of the right-hand side of the equation $\ln y=a\ln x$ is

$$(a\ln x)'=a\cdot\frac{1}{x}.$$

Consequently,

$$\frac{1}{y}\cdot y'=\frac{a}{x},$$

whence

$$y'=(x^a)'=y\cdot\frac{a}{x}=x^a\cdot\frac{a}{x}=ax^{a-1}. \qquad \text{(IX)}$$

Now let $x<0$. We put $x=-z$ $(z>0)$. By the rule for differentiating a function of a function and formula (IX), we find

$$(x^a)'=[(-z)^a]'=[(-1)^a z^a]'=(-1)^a\cdot az^{a-1}\cdot z'.$$

Since $z=-x$, we have $z'=-1$, whence

$$(x^a)'=(-1)^a a\cdot z^{a-1}(-1)=(-1)^{a+1}az^{a-1}$$
$$=(-1)^{a-1}(-1)^2 az^{a-1}.$$

or

$$(x^a)'=(-1)^{a-1}az^{a-1}=a(-z)^{a-1},$$

i.e.

$$(x^a)'=ax^{a-1}.$$

§ 58. **Derivative of exponential functions.** We take natural logarithms of both sides of the equation $y=a^x$ $(a>0)$ to give

$$\ln y=x\ln a.$$

Here $\ln y$ is a function of a function, since the logarithm is taken of y, whilst y is a function of x. Hence we find

$$\frac{1}{y}y'=\ln a\ *,$$

so that

$$y'=(a^x)'=y\ln a=a^x\ln a. \qquad \text{(XVI)}$$

Thus *the derivative of the exponential function a^x is equal to the exponential function itself multiplied by the natural logarithm of its base.*

In the particular casy of $a=e$, we have the function $y=e^x$. We obtain from formula (XVI):

$$y'=(e^x)'=e^x\ln e=e^x, \qquad \text{(XVII)}$$

i.e. *the derivative of the exponential function e^x is equal to the function itself.*

Example. Differentiate the function

$$y=7^{4x^2}.$$

Solution.

$$\left.\begin{aligned} y'&=7^{4x^2}\cdot\ln 7\cdot(4x^2)' \\ &=7^{4x^2}\ln 7\cdot 8x=8\ln 7\cdot x\cdot 7^{4x^2} \end{aligned}\right\}\ \begin{array}{l}\text{(by formulae}\\ \text{XVI and VI).}\end{array}$$

---

* On the assumption that the derivative of $y=a^x$ exists, cf. remark at the beginning of § 57.

§ 59. Derivatives of the inverse trigonometric functions *.

1. *Derivative of the function arc sin x.* Let

$$y = \arc \sin x \quad (-1 < x < 1),$$

where $\left(-\frac{1}{2}\pi < y < +\frac{1}{2}\pi\right)$. Then $\sin y = x$.

Here $\sin y$ is a function of a function, since y is a function of x. On differentiating both sides of the equation $\sin y = x$ with respect to x we find that

$$\cos y \cdot y' = 1,$$

whence
$$y' = \frac{1}{\cos y}.$$

Since $\sin y = x$, $\cos y = \sqrt{1-x^2}$.** Thus

$$y' = (\operatorname{arc} \sin x)' = \frac{1}{\sqrt{1-x^2}}. \qquad \text{(XVIII)}$$

2. *Derivative of the function* arc cos x. We write

$$y = \operatorname{arc} \cos x \quad (-1 < x < 1),$$

where $0 < y < \pi$; then $\cos y = x$.

On differentiating both sides of this equation with respect to x we find that

$$-\sin y \cdot y' = 1,$$

whence

$$y' = -\frac{1}{\sin y}.$$

---

* These formulae are deduced on the assumption that the derivatives of the inverse trigonometric functions exist, cf. remark at the beginning of § 57.

** We take the plus sign in front of the square root because we only consider the values of $y = \operatorname{arc} \sin x$ lying between $-\frac{1}{2}\pi$ and $+\frac{1}{2}\pi$ $\left(-\frac{1}{2}\pi < y < +\frac{1}{2}\pi\right)$ and $\cos y$ is positive between these limits.

Since $\cos y = x$, $\sin y = \sqrt{1-x^2}$.* Thus

$$y' = (\text{arc cos } x)' = -\frac{1}{\sqrt{1-x^2}}. \tag{XIX}$$

3. *Derivative of the function* arc tan x. We write

$$y = \text{arc tan } x \quad (-\infty < x < \infty),$$

where $-\frac{1}{2}\pi < y < \frac{1}{2}\pi$. Then $\tan y = x$.

On differentiating both sides of this equation we find that

$$\frac{1}{\cos^2 y} \cdot y' = 1,$$

whence

$$y' = \cos^2 y.$$

Since $\tan y = x$ and $\cos^2 y = \frac{1}{1+\tan^2 y}$, we have $\cos^2 y = \frac{1}{1+x^2}$,

whence

$$y' = (\text{arc tan } x)' = \frac{1}{1+x^2}. \tag{XX}$$

4. *Derivative of the function arc cot x.* We write

$$y = \text{arc cot } x \quad (-\infty < x < \infty),$$

where $0 < y < \pi$. Then $\cot y = x$.

Differentiating both sides of this equation with respect to x, we obtain

$$-\frac{1}{\sin^2 y} \cdot y' = 1,$$

whence
$$y' = -\sin^2 y.$$

* We take the plus sign in front of the square root because we only consider the values of $y = \text{arc cos } x$ between 0 and π $(0 < y < \pi)$ and $\sin y$ is positive between these limits.

Since $\cot y = x$ and $\sin^2 y = \frac{1}{1+\cot^2 y}$, we have $\sin^2 y = \frac{1}{1+x^2}$, whence $y' = (\text{arc cot } x)' = -\frac{1}{1+x^2}$. (XXI)

Examples. Find the derivatives of the following functions:

1. $y = \text{arc tan } 3x^2$.

Solution.

$$y' = \frac{1}{1+9x^4} \cdot (3x^2)' = \frac{6x}{1+9x^4} \quad \text{(by formulae XX and VI).}$$

2. $y = \text{arc sin } (3x - 4x^3)$.

Solution.

$$\left.\begin{aligned} y' &= \frac{1}{\sqrt{1-(3x-4x^3)^2}}(3x-4x^3)' \\ &= \frac{3-12x^2}{\sqrt{1-9x^2+24x^4-16x^6}} = \frac{3}{\sqrt{1-x^2}} \end{aligned}\right\} \text{(by formulae XVIII and VI)}$$

3. $y = \text{arc cot } \frac{x+a}{1-ax}$.

Solution.

$$\left.\begin{aligned} y' &= -\frac{1}{1+\left(\frac{x+a}{1-ax}\right)^2} \cdot \left(\frac{x+a}{1-ax}\right)' \\ &= -\frac{(1-ax)^2}{(1+a^2)(1+x^2)} \cdot \frac{1+a^2}{(1-ax)^2} = -\frac{1}{1+x^2} \end{aligned}\right\} \text{(by formulae XXI and VI).}$$

EXERCISES

On §§ 45—52.

Find the derivatives of the following functions:

1. $y = 5x^3 - 3x^2 + 6$. Ans. $y' = 15x^2 - 6x$.

2. $y=x^5-\frac{1}{4}x^4+3\cdot 5x^2-x.$ Ans. $y'=5x^4-x^3+7x-1.$

3. $y=x^{a+b}.$ Ans. $y'=(a+b)\,x^{a+b-1}.$

4. $y=ax^2+bx+c.$ Ans. $y'=2ax+b.$

5. $y=\frac{2}{x}+\frac{x}{2}.$ Ans. $y'=-\frac{2}{x^2}+\frac{1}{2}.$

6. $y=\frac{2}{x^2}+\frac{3}{x^3}.$ Ans. $y'=-\frac{4}{x^3}-\frac{9}{x^4}.$

7. $y=2\sqrt{x}-\sqrt[3]{x}.$ Ans. $y'=\frac{1}{\sqrt{x}}-\frac{1}{3\sqrt[3]{x^2}},$

8. $y=0{\cdot}4\sqrt[4]{x}+\sqrt[3]{3}.$ Ans. $y'=\frac{0{\cdot}1}{\sqrt[4]{x^3}}.$

9. $y=6x^{\frac{7}{2}}+4x^{\frac{5}{2}}+2x^{\frac{3}{2}}.$ Ans. $y'=21x^{\frac{5}{2}}+10x^{\frac{3}{2}}+3x^{\frac{1}{2}}.$

10. $y=x^{-2}-4x^{-\frac{1}{2}}.$ Ans. $y'=-\frac{2}{x^3}+\frac{2}{\sqrt{x^3}}.$

11. $y=\sqrt{3x}+\sqrt[3]{x}+\frac{1}{x}.$ Ans. $y'=\frac{\sqrt{3}}{2\sqrt{x}}+\frac{1}{3\sqrt[3]{x^2}}-\frac{1}{x^2}.$

12. $y=\frac{px^3}{\sqrt{x}}+\frac{q\sqrt{x}}{\sqrt[3]{x}}-\frac{r}{\sqrt[5]{x}}.$ Ans. $y'=\frac{5p}{2}\sqrt{x^3}+\frac{q}{6\sqrt[6]{x^5}}+$

$$+\frac{r}{5x\sqrt[5]{x}}.$$

13. $y=t^{1\cdot 2}-\frac{3}{t^{1\cdot 5}}.$ Ans. $y'=1\cdot 2t^{0\cdot 2}+\frac{4\cdot 5}{t^{2\cdot 5}}.$

14. $y=\frac{mz^2+nz+2p}{g+r}.$ Ans. $y'=\frac{2mz+n}{g+r}.$

15. $y=\frac{x^3+x^2+x+1}{3x}.$ Ans. $y'=\frac{1}{3}\left(2x+1-\frac{1}{x^2}\right).$

16. $s=\sqrt{t}\,(2t^3-\sqrt{t}+1)$. Ans. $s'=7t^2\sqrt{t}-1+\dfrac{1}{2\sqrt{t}}$.

17. $f(x)=x^2-5x+6$. Find: $f(1)$, $f'(1)$, $f(2)$, $f'(2)$, $f(3)$, $f'(3)$, $f(\frac{1}{2})$, $f'(\frac{1}{2})$, $f(a)$, $f'(a)$.

Ans. 2; -3; 0; -1; 0; 1; 3·75; -4; a^2-5a+6; $2a-5$.

18. $f(x)=x^5-3x^3+1$. Show that $f'(a)=f'(-a)$.

19. Find the values of x and y for which the derivatives of the following functions vanish:

a) $y=2x^3-9x^2$, b) $y=2x^3+3x^2-12x-18$,

c) $y=\dfrac{1}{4}x^4-2x^2+\dfrac{1}{4}$, d) $y=x^4-2x^3+4$.

Ans. a) $(0, 0)$; $(3, -27)$; b) $(-2, 2)$; $(1, -25)$; c) $\left(-2, -3\dfrac{3}{4}\right)$; $\left(0, \dfrac{1}{4}\right)$; $\left(2, -3\dfrac{3}{4}\right)$; d) $(0; 4)$; $\left(\dfrac{3}{2}, \dfrac{37}{16}\right)$.

20. Find the equation of the tangent to the curve $y=4x^2+4x-3$ at the point whose abscissa is -1.

Ans. $4x+y+7=0$.

21. Find the equation of the tangent to the curve $y=2-4x-x^2$ at the point whose abscissa is -3.

Ans. $2x-y+11=0$.

22. Find the equations of the tangents to the curve $y=x^3+x^2$ whose slopes are equal to 8.

Ans. $8x-y+12=0$; $216x-27y-176=0$.

23. Find the equations of the tangents to the curve $y=2x^3+4x^2-x$ whose slopes are equal to $\dfrac{1}{2}$.

Ans. $x-2y+9=0$; $27x-54y-7=0$.

24. Given the curve $y=x^3-2x^2+x-2$, find the equations of its tangents that are parallel to the straight line $7x-4y+28=0$.

Ans. $7x-4y-17=0$; $189x-108y-209=0$.

25. Find the points on the curve $y=x^3-x^2+2x+3$ at which the tangents are parallel to the straight line $3x-y-7=0$.

Ans. $(1, 5)$; $\left(-\frac{1}{3}, \frac{59}{27}\right)$.

26. Find the angle between the tangents to the curve $y=x^3-3x^2+4x-12$ at the points with abscissae -1 and 1.

Ans. $\arctan\frac{6}{7}$.

27. Find the area of the triangle formed by the co-ordinate axes and the tangent to the curve $y=x^3$ at the point $(3, 27)$.

Ans. 54 sq. units.

28. Find the equation of the tangent to the parabola $y=x^2-4x+7$ which is perpendicular to the straight line joining the origin to the vertex of the parabola.

Ans. $6x+9y-38=0$.

29. Find the equation of the tangent to the curve $y=+\sqrt{x+6}$ at the point of intersection of the curve with the line $y=x$.

Ans. $x-6y+45=0$.

30. The distance s in metres traversed by a particle after t sec is given by the equation $s=2t^3-3$. Find the speed of the particle: (*a*) 1 sec, (*b*) 2 sec after the initial instant.

Ans. (*a*) 6 m/sec, (*b*) 24 m/sec.

31. If a particle moves in accordance with the law $s=t^2-4t+5$ when is its velocity zero?

Ans. $t=2$ sec.

32. The height s in metres reached after t sec by a body thrown vertically upwards with an initial velocity of v_0 m/sec is given by the formula

$$s=v_0t-4{\cdot}9t^2 .$$

(*a*) Find the velocity of the body at any instant t.
Putting $v_0=100$ m/sec, find:
(*b*) the velocity at the end of the 2nd second,
(*c*) at the end of the 15th second,
(*d*) after how many seconds the body attains its greatest height.

Ans. (*a*) $v=v_0-9{\cdot}8\,t$, (*b*) 80·4 m/sec, (*c*) —47 m/sec, (*d*) approximately 10·2 sec.

Find the derivatives of the following functions:

33. $y=(2x+3)(x^2+3x-1)$. Ans. $y'=6x^2+18x+7$.

34. $y=(1+4x^3)(1+2x^2)$. Ans. $y'=4x(1+3x+10x^3)$.

35. $y=(x^2+4x-3)(3x^2+12x+12)$.

Ans. $y'=6(x+2)(2x^2+8x+1)$.

36. $y=x(2x-1)(3x+2)$. Ans. $y'=2(9x^2+x-1)$.

37. $s=(\sqrt{t}+1)\left(1-\dfrac{1}{\sqrt{t}}\right)$. Ans. $s'=\dfrac{1+t}{2t\sqrt{t}}$.

38. $f(x)=(1+x^2)\left(3-\dfrac{1}{x^3}\right)$. Ans. $f'(x)=6x+\dfrac{1}{x^2}+\dfrac{3}{x^4}$.

Find $f'(1)$ and $f'(-)$. Ans. 10 ; —2.
Find the derivatives of the following functions:

39. $y=\dfrac{x+a}{x-a}$. Ans. $y'=-\dfrac{2a}{(x-a)^2}$.

40. $y=\dfrac{x^2-4}{x^2+4}$. Ans. $y'=\dfrac{16x}{(x^2+4)^2}$.

41. $y=\dfrac{x^3}{4-x}$. Ans. $y'=\dfrac{2x^2(6-x)}{(4-x)^2}$.

42. $v = \dfrac{u^2+u-2}{u^3-1}$. Ans. $v' = -\dfrac{u^2+4u+1}{(u^2+u+1)^2}$.

43. $y = \dfrac{1-\sqrt{x}}{1+\sqrt{2x}}$. Ans. $y' = -\dfrac{1+\sqrt{2}}{2\sqrt{x}\,(1+\sqrt{2x})^2}$.

44. $s = \dfrac{1-\sqrt[3]{t^2}}{1+\sqrt[3]{t^2}}$. Ans. $s' = -\dfrac{4}{3\sqrt[3]{t}\,(1+\sqrt[3]{t^2})^2}$.

45. $f(x) = \dfrac{a-x}{a+x}$. Find $f'(a)$. Ans. $-\dfrac{1}{2a}$.

46. What is the angle formed by the tangent to the curve $y = \dfrac{x}{1+x^2}$ at the origin with the Ox axis?

Ans. $\dfrac{\pi}{4}$.

Find the derivatives of the following functions:

47. $y = (x^3-1)^{100}$. Ans. $y' = 300x^2(x^3-1)^{99}$.

48. $y = \sqrt{1+x^2}$. Ans. $y' = \dfrac{x}{\sqrt{1+x^2}}$.

49. $y = \sqrt[3]{1+x^2}$. Ans. $y' = \dfrac{2x}{3\sqrt[3]{(1+x^2)^2}}$.

50. $y = \sqrt{x+\sqrt{x}}$. Ans. $y' = \dfrac{2\sqrt{x}+1}{4\sqrt{x}\sqrt{x+\sqrt{x}}}$.

51. $y = (x-1)\sqrt{x^2+1}$. Ans. $y' = \dfrac{2x^2-x+1}{\sqrt{x^2+1}}$.

52. $y = (3x-1)^2(x-1)^3$. Ans. $y' = 3(5x-3)(3x-1)$
$(x-1)^2$.

53. $y=\dfrac{2x-1}{\sqrt{x^2+1}}$. Ans. $y'=\dfrac{x+2}{\sqrt{(x^2+1)^3}}$.

54. $y=\sqrt{\dfrac{x+1}{x-1}}$. Ans. $y'=\dfrac{1}{(1-x)\sqrt{x^2-1}}$.

55. $y=\dfrac{x}{\sqrt{a^2+x^2}}$. Ans. $y'=\dfrac{a^2}{(a^2+x^2)\sqrt{a^2+x^2}}$.

56. $y=\dfrac{\sqrt{a^2+x^2}}{x}$. Ans. $y'=-\dfrac{a^2}{x^2\sqrt{a^2+x^2}}$.

57. $y=\dfrac{1}{x+\sqrt{x^2-1}}$. Ans. $y'=1-\dfrac{x}{\sqrt{x^2-1}}$.

58. $y=\dfrac{x}{x+\sqrt{a^2+x^2}}$. Ans. $y'=\dfrac{2x^2+a^2}{a^2\sqrt{a^2+x^2}}-\dfrac{2x}{a^2}$.

59. $y=\dfrac{1}{x-\sqrt{a^2+x^2}}$. Ans. $y'=-\dfrac{x+\sqrt{a^2+x^2}}{a^2\sqrt{a^2+x^2}}$.

On § 54.

Find the derivatives of the following functions:

60. $y=\cos 5x$. Ans. $y'=-5\sin 5x$.

61. $y=\sin(3ax)$. Ans. $y'=3a\cos(3ax)$.

62. $y=\cos\dfrac{a}{x}$. Ans. $y'=\dfrac{a}{x^2}\sin\dfrac{a}{x}$.

63. $y=\tan(2x+3)$. Ans. $y'=\dfrac{2}{\cos^2(2x+3)}$.

64. $y=\tan x-x$. Ans. $y'=\tan^2 x$.

65. $s=\sin\sqrt{1-t^2}$. Ans. $s'=-\dfrac{t}{\sqrt{1-t^2}}\cos\sqrt{1-t^2}$.

66. $v=\cos\dfrac{1-\sqrt{u}}{1+\sqrt{u}}$. Ans. $v'=\dfrac{1}{\sqrt{u}(1+\sqrt{u})^2}\sin\dfrac{1-\sqrt{u}}{1+\sqrt{u}}$.

67. $y=\cos^3 x^2$. Ans. $y'=-6x\sin x^2\cos^2 x^2$.

68. $y=-\dfrac{3}{5}\cot^5\dfrac{x}{3}+\cot^3\dfrac{x}{3}-3\cot\dfrac{x}{3}-x$. Ans. $y'=\cot^6\dfrac{x}{3}$.

69. $y=\sqrt{1+\tan\left(x+\dfrac{1}{x}\right)}$.

Ans. $y'=\dfrac{x^2-1}{2x^2\sqrt{1+\tan\left(x+\dfrac{1}{x}\right)}\cos^2\left(x+\dfrac{1}{x}\right)}$.

70. $f(u)=\sin^2 u$. Find $f'\left(\dfrac{\pi}{4}\right)$. Ans. 1.

71. $f(x)=\dfrac{1-\cos x}{1+\cos x}$. Find $f'\left(\dfrac{1}{2}\pi\right)$, $f'(0)$. Ans. 2, 0.

72. What are the angles of intersection with the Ox axis of the tangents to $y=\sin x$ at the points $x=0$, $x=\pi$? And at what points in the interval $-\pi\leqq x\leqq+\pi$ are the tangents parallel to Ox?

Ans. $\dfrac{\pi}{4}$; $\dfrac{3\pi}{4}$; $-\dfrac{\pi}{2}$; $+\dfrac{\pi}{2}$.

On §§ 56—58.

Find the derivatives of the following functions:

73. $y=\ln(x-2)$. Ans. $y'=\dfrac{1}{x-2}$.

74. $y=\ln(ax+b)$. Ans. $y'=\dfrac{a}{ax+b}$.

75. $y=\ln(x^2+2x)$. Ans. $y'=\dfrac{2(x+1)}{x^2+2x}$.

76. $y=\log_5(3x^2+1)$. Ans. $y'=\dfrac{6x\log_5 e}{3x^2+1}$.

77. $y=\log_a(x+x^3)$. Ans. $y'=\dfrac{1+3x^2}{x+x^3}\log_a e$.

78. $y=\log_3(x+\sqrt{x})$. Ans. $y'=\dfrac{2\sqrt{x}+1}{2(x\sqrt{x}+x)}\log_3 e$.

79. $y=\ln\ln x$. Ans. $y'=\dfrac{1}{x\ln x}$.

80. $y=\ln(1+\ln x)$. Ans. $y'=\dfrac{1}{x(1+\ln x)}$.

81. $y=\ln x^2$. Ans. $y'=\dfrac{2}{x}$.

82. $y=\ln^2 x$. Ans. $y'=\dfrac{2\ln x}{x}$.

83. $y\ \sqrt{1+\ln^2 x}$. Ans. $y'=\dfrac{\ln x}{x\sqrt{1+\ln^2 x}}$.

84. $y=x\ln x$. Ans. $y'=\ln x+1$.

85. $y=\ln\dfrac{a+x}{a-x}$. Ans. $y'=\dfrac{2a}{a^2-x^2}$.

86. $y=\ln(2x^2-4x+3)$. Ans. $y'=\dfrac{4(x-1)}{2x^2-4x+3}$.

87. $y=\ln\sqrt{x^2+4x+3}$. Ans. $y'=\dfrac{x+2}{x^2+4x+3}$.

88. $s=\dfrac{1}{2\sqrt{3}}\ln\dfrac{3t-\sqrt{3}}{3t+\sqrt{3}}$. Ans. $s'=\dfrac{1}{3t^2-1}$.

89. $v=\mathrm{n}\dfrac{1}{\sqrt{3-4u+u^2}}$. Ans. $v'=\dfrac{2-u}{3-4u+u^2}$.

90. $y=\ln(x+\sqrt{1+x^2})$. Ans. $y'=\dfrac{1}{\sqrt{1+x^2}}$.

91. $y=\ln(3x+\sqrt{9x^2+2})$. Ans. $y'=\dfrac{3}{\sqrt{9x^2+2}}$.

92. $y=\dfrac{1}{3}\ln(x^3+\sqrt{x^6-a^6})$. Ans. $y'=\dfrac{x^2}{\sqrt{x^6-a^6}}$.

93. $s=\dfrac{1}{a}\ln\dfrac{\sqrt{a^2+t^2}-a}{t}$. Ans. $s'=\dfrac{1}{t\sqrt{a^2+t^2}}$.

94. $y=\ln\sqrt{x^2+4}+\dfrac{2}{x^2+4}$. Ans. $y'=\dfrac{x^3}{(x^2+4)^2}$.

95. $f(x)=\sqrt[3]{\ln x}$. Find $f'(e)$. Ans. $\dfrac{1}{3e}$.

96. $y=e^{4x-3}$. Ans. $y'=4e^{4x-3}$.

97. $y=a^{3x}$. Ans. $y'=3a^{3x}\ln a$.

98. $s=7^{t^3+3t}$. Ans. $s'=3(t^2+1)\,7^{t^3+3t}\ln 7$.

99. $p=\dfrac{1}{5^q}$. Ans. $p'=\dfrac{-\ln 5}{5^q}$.

100. $y=e^{-\frac{1}{x}}$. Ans. $y'=e^{-\frac{1}{x}}\cdot\dfrac{1}{x^2}$.

101. $y=4^{1-\sqrt{x}}$. Ans. $y'=-\dfrac{1}{2\sqrt{x}}\cdot 4^{1-\sqrt{x}}\ln 4$.

102. $y=e^{e^x}$. Ans. $y'=e^x\cdot e^{e^x}$.

103. $y=2x-a\ln(e^{\frac{x}{a}}+1)$. Ans. $y'=\dfrac{e^{\frac{x}{a}}+2}{e^{\frac{x}{a}}+1}$.

104. $y=a^x e^x$. Ans. $y'=a^x e^x(1+\ln a)$.

105. $y=\dfrac{a^{b+cx}e^{b+cx}}{c(1+\ln a)}$. Ans. $y'=a^{b+cx}e^{b+cx}$.

106. $y=\frac{a}{2}\left(e^{\frac{x}{a}}+e^{-\frac{x}{a}}\right)$. Ans. $y'=\frac{1}{2}\left(e^{\frac{x}{a}}-e^{-\frac{x}{a}}\right)$.

107. $y=\frac{e^x-1}{e^x+1}$. Ans. $y'=\frac{2e^x}{(e^x+1)^2}$.

108. $y=\frac{e^x-e^{-x}}{e^x+e^{-x}}$. Ans. $y'=\frac{4}{(e^x+e^{-x})^2}$.

109. $y=\ln\tan\left(\frac{\pi}{4}+\frac{1}{2}x\right)$. Ans. $y'=\sec x$.

110. $y=\ln\cos x$. Ans. $y'=-\tan x$.

111. $y=\sin\ln x$. Ans. $y'=\frac{\cos\ln x}{x}$.

112. $y=\ln\sqrt{\frac{1+\sin x}{1-\sin x}}$. Ans. $y'=\frac{1}{\cos x}$.

113. $y=e^{\sin x}$. Ans. $y'=\cos x e^{\sin x}$.

114. $y=a^{\tan nx}$. Ans. $y'=\frac{na^{\tan nx}\ln a}{\cos^2 nx}$.

On § 59.

Find the derivatives of the following functions:

115. $y=\arcsin\frac{x}{a}$. Ans. $y'=\frac{1}{\sqrt{a^2-x^2}}$.

116. $y=\arctan\frac{2x}{1-x^2}$. Ans. $y'=\frac{2}{1+x^2}$.

117. $y=\arctan\sqrt{x^2+2x}$. Ans. $y'=\frac{1}{(x+1)\sqrt{x^2+2x}}$.

118. $y=\operatorname{arc\,cot}\frac{x}{\sqrt{a^2-x^2}}$. Ans. $y'=-\frac{1}{\sqrt{a^2-x^2}}$.

119. $y=\arccos\frac{x^2-a^2}{x^2+a^2}$. Ans. $y'=-\frac{2a}{x^2+a^2}$.

120. $y = \operatorname{arc\,cot} \sqrt{\dfrac{1-x}{x}}$. Ans. $y' = \dfrac{1}{2\sqrt{x-x^2}}$.

121. $y = \arctan \frac{1}{2}(e^x - e^{-x})$. Ans. $y' = \dfrac{2}{e^x + e^{-x}}$.

122. $y = \arccos \dfrac{e^x - e^{-x}}{e^x + e^{-x}}$. Ans. $y = -\dfrac{2}{e^x + e^{-x}}$.

123. $y = \sqrt{1-x^2} \arcsin x - x$. Ans. $y' = -\dfrac{x \arcsin x}{\sqrt{1-x^2}}$.

124. $y = x\sqrt{a^2 - x^2} + a^2 \arcsin \dfrac{x}{a}$. Ans. $y' = 2\sqrt{a^2 - x^2}$.

125. $y = x \arcsin \sqrt{1-x^2} - \sqrt{1-x^2}$. Ans. $y' = \arcsin \sqrt{1-x^2}$.

CHAPTER 7

THE INVESTIGATION OF FUNCTIONS WITH THE AID OF THE DERIVATIVE

§ 60. Variation of a Function. We established in § 36 that a function may in general be represented geometrically by means of a curve. Let fig. 52 be the graph of a function $y=f(x)$ given in the

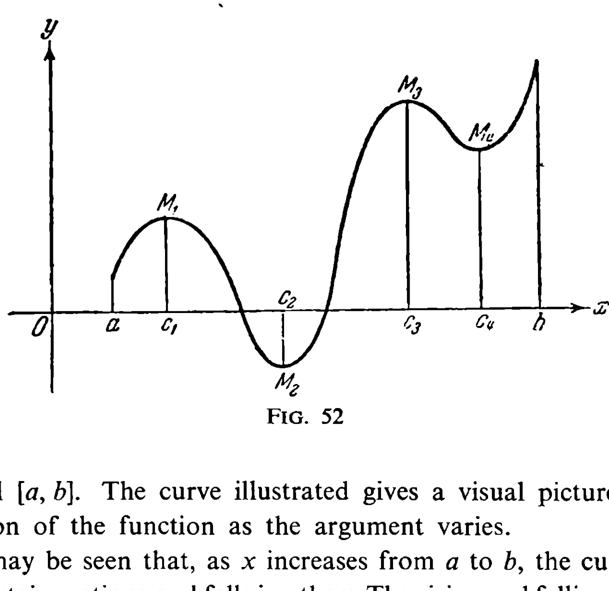


FIG. 52

interval $[a, b]$. The curve illustrated gives a visual picture of the variation of the function as the argument varies.

It may be seen that, as x increases from a to b, the curve rises over certain sections and falls in others. The rising and falling sections correspond to the intervals in which the function is increasing and decreasing. Fig. 52 shows visually that the function is increasing in the intervals (a, c_1), (c_2, c_3), (c_4, b) and decreasing in the intervals (c_1, c_2), (c_3, c_4). Points M_1, M_3 are the points at which the curve rises highest as compared with neighbouring portions, whilst M_2, M_4 are the points at which it falls lowest as compared with neighbouring portions.

For the values of the argument $x=c_1$ and $x=c_3$ (fig. 52) corresponding to the points of highest rise of the graph, the function $y=f(x)$ has its greatest values as compared with its values at neighbouring points of the interval $[a, b]$; at the values of the argument where the graph falls lowest ($x=c_2$ and $x=c_4$ in fig. 52), the function $y=f(x)$ has its least values as compared with its values at neighbouring points.

The rise and fall of a function, and the maximum and minimum values which it attains throughout the interval in which it is defined, represent characteristic elements as regards the variation of the function. We consider in the present chapter how these characteristic features may be determined for the variation of any given function. It will be seen that the problem is solved with the aid of derivatives.

§ 61. Increase and decrease of a function in an interval.

DEFINITION. *If the value of the function $y=f(x)$ increases as the argument x increases over the interval from a to b $(a<b)$, the function is said to be* increasing *in this interval.*

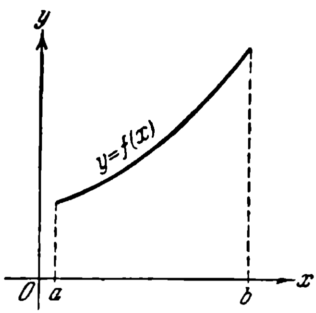


FIG. 53

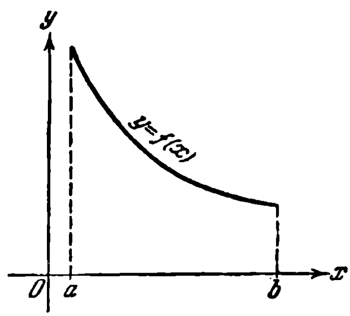


FIG. 54

Similarly, if the value of $y=f(x)$ decreases as the argument x increases over the interval from a to b, the function is said to be *decreasing.*

It follows from the definitions that the graph of a function which is increasing in the interval (a, b) consists of a curve which rises as the abscissae increase (fig. 53).

The graph of a function which is decreasing in the interval (a, b) consists of a curve which falls as the abscissae increase (fig. 54).

To trace the variation of a function, we want to know how to discover the intervals in which the function is increasing and those in which it is decreasing. We shall establish an analytical test for this purpose. The test is expressed in the form of the following theorems:

THEOREM (sufficient test for increasing function). *If the derivative of a function is positive for all values of x in the interval (a, b), the function is increasing in this interval.*

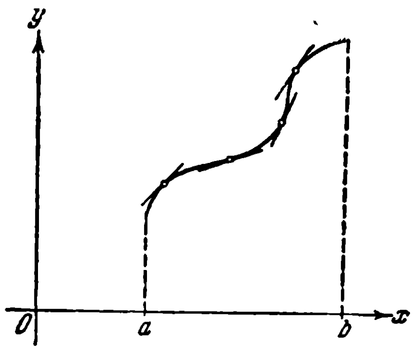


FIG. 55

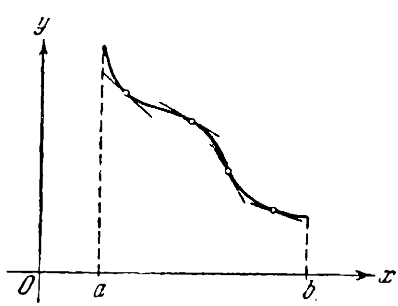


FIG. 56

Whilst a rigorous proof is outside the scope of the present course, the validity of the theorem may readily be seen from geometrical considerations. For the derivative is the slope of the tangent to the graph of the function or more simply, the slope of the graph. Hence if the derivative is positive, the tangent, forming an acute angle with Ox, retains an upward slope in the interval (a, b), whilst the curve itself also moves upwards (fig. 55).

THEOREM (sufficient test for decreasing function). *If the derivative of a function is negative for all values of x in the interval (a, b), the function is decreasing in this interval.*

For, if the derivative is negative in (a, b), the tangent to the graph of the function, forming an obtuse angle with Ox, retains a downward slope, whilst the curve itself also travels downwards (fig. 56).

Example. Find the intervals in which the function

$$y = x^3 - 3x^2 + 5$$

is increasing and decreasing.

Solution. We find the derivative of the function as

$$y' = 3x^2 - 6x.$$

We now have to find the values of x for which the derivative is positive, and those for which it is negative. We do this by splitting $3x^2 - 6x$ into factors and obtain

$$y' = 3x(x-2).$$

It now becomes evident that the product $3x(x-2)$ is positive for all $x < 0$, since x and $x-2$ both have the same sign. With x positive but less than two the derivative is negative, whilst with $x > 2$ it again becomes positive. The function is therefore increasing in

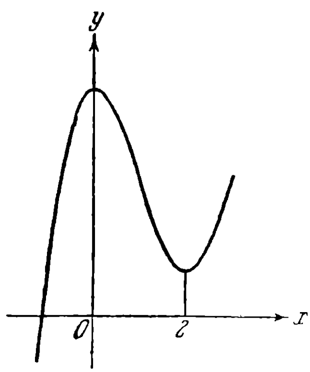


FIG. 57

the interval from $-\infty$ to 0, decreasing in the interval from 0 to 2, and increasing in the interval from 2 to $+\infty$. The graph of the function is shown in fig. 57.

§ 62. Maxima and minima of a function. To find the extrema of a function. 1. We have solved in the previous section the problem of finding the intervals in which a function is increasing or decreas-

ing. We now turn to the problem of finding the points (values of the argument (x) at which a function $f(x)$ attains its greatest and least values *as compared with its values at neighbouring points* (cf. § 60).

DEFINITION. *The function $f(x)$ has a maximum [minimum] at the point $x=c$ if there is a neighbourhood of this point such that $f(x)<f(c)$ [($f(x)>f(c)$] for all values of x in this neighbourhood.*

Figure 52 gives a visual illustration of this definition. It is clear, for instance, that the function whose graph is shown has a maximum at the point c_1, for we have $f(c_1)>f(x)$ for all the $x \neq c_1$ contained in the interval (a, c_2). The function has another maximum at the point c_3, whilst it has minima at c_2 and c_4.

It must be observed that the values of the function at the points where it has maxima or minima are not necessarily the greatest or least values of the function in the interval $[a, b]$. Fig. 52 shows, in fact, that the function attains its greatest value at $x=b$, i.e. at the end of the interval, and not at the points where it has maxima. In addition to all this, though the function has a minimum at the point $x=c_4$, its value $f(c_4)$ at this point is greater than its value $f(c_1)$ at the point $x=c_1$ where it has a maximum. The maxima and minima referred to in our definition are thus often termed *relative* maxima and minima; the function has its greatest and least values at points of maxima and minima only by comparison with its neighbouring values.

The terms "maximum" and "minimum" are covered by the common name "extremum".

2. We now establish a rule to enable us to find the extrema (i.e. maxima and minima) of a function. This rule is based on a property of functions continuous in an interval, which may be expressed in the form of the following theorem (the rigorous proof will be omitted).

THEOREM. *If the function $f(x)$, continuous in the closed interval $[a, b]$, has values of differing signs at the ends of the interval, there must be at least one interior point $x=x_1$ of the interval at which the function vanishes $[f(x_1)=0]$.*

Fig. 58 gives a visual illustration of the truth of this theorem: the continuous curve representing the graph of a function contin-

uous in the interval $[a, b]$ must cross the Ox axis at least once in the interval in order to pass from one side of the axis to the other.

We now consider the function $y=f(x)$, continuous in the interval (a, b) and having a continuous derivative $f'(x)$ in (a, b).

Suppose that the derivative $f'(x)$ does not vanish for any x in the interval (a, b). Then it must preserve a fixed sign throughout

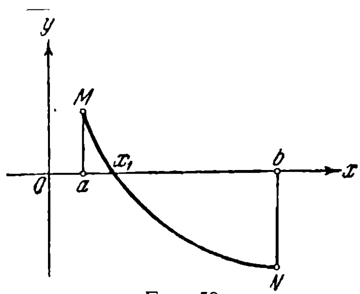


FIG. 58

the interval. For suppose the derivative be positive for $x=x_1$ and negative for $x=x_2$; then it must vanish, by virtue of its continuity and the theorem just proved, for at least one value $x+x_0$ between x_1 and x_2. But this is impossible, since x_0 lies inside the interval (a, b) and we have assumed that the derivative does not vanish throughout (a, b). If the derivative $f'(x)$ preserves the same sign throughout (a, b), the function $y=f(x)$ is either always increasing or always decreasing in the interval (cf. § 61), and no extremum can exist.

Hence, if $f(x)$ has an extremum in the interval (a, b), it must be at points where the derivative $f'(x)$ vanishes.

Now let $f(x)$ vanish in the interval (a, b), but only at a *finite* number of points, say $c_1 < c_2 < c_3 \ldots < c_k$. Now, by what has been proved, the derivative $f'(x)$ preserves a fixed sign in each of the intervals (a, c_1), (c_1, c_2), ..., (c_k, b). We take any one of the points $c_1, c_2, \ldots, c_k$, say c_1.

Since $f(x)$ is continuous in (a, b) it has a definite value $f(c_1)$ at the point c_1. Now let the derivative be positive in the interval (a, c_1) and negative in (c_1, c_2). Then the function $f(x)$ is increasing in the first interval and decreasing in the second; in other words, $f(x)$ changes over from increasing to decreasing at the point c_1, and its value $f(c_1)$ at this point must be the greatest of its values

in the intervals (a, c_1) and (c_1, c_2), i.e. $f(x)$ has a maximum at $x=c_1$. If, on the contrary, $f'(x)$ is negative in the interval (a, c_1) and positive in (c_1, c_2), the function $f(x)$ must be decreasing in the first interval and increasing in the second, and $f(x)$ therefore has a minimum at $x=c_1$. Finally, suppose $f'(x)$ has the same sign in the intervals (a, c_1) and (c_1, c_2), say $f'(x)>0$. Then $f(x)$ must be an increasing function in both intervals, i.e. $f(x)$ is increasing throughout (a, c_2), in which case it has neither a maximum nor a minimum at $x=c_1$. Similarly, if $f'(x)<0$ in both the intervals (a, c_1) and (c_1, c_2), $f(x)$ cannot have an extremum at $x=c_1$.

The same arguments enable us to solve our problem regarding the extrema at each of the remaining points

$$c_2, c_3, \ldots, c_k.$$

The results obtained permit of a very simple geometrical interpretation.

The points on the graph of $f(x)$ corresponding to its extrema are points at which the tangent to the curve is parallel to the Ox

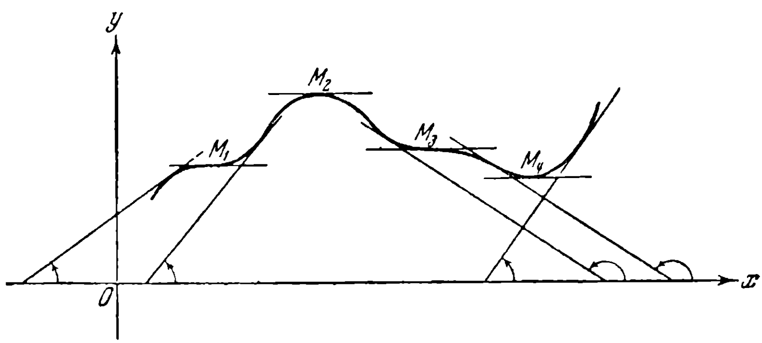


FIG. 59

axis (points M_2 and M_4 on fig. 59). Thus at the points where the function has an extremum, the derivative

$$f'(x)=0.$$

Conversely, if $f'(x)$ vanishes at some value of x, the tangent at the corresponding point of the curve is parallel to Ox. If the derivative does not change sign on passing through this value of x,

the angle of inclination of the tangent remains either obtuse or acute to the left and the right of the point of the curve in question. The curve is in this case either always rising or always falling, and the point where the tangent is parallel to Ox is clearly not one at which $f(x)$ has an extremum, but is, in fact, a so-called "point of inflexion" (points M_1 and M_3 of the curve in fig. 59). Since a change in the sign of the derivative on passing through the value of x in question means that the curve starts to fall instead of rise, or rise instead of fall, at this point, the function has an extremum in this case (cf. points M_2 and M_4 on fig. 59).

3. The arguments of 2. above lead to the following rule for finding the values of x in the interval (a, b) at which the function $f(x)$ has extrema.

For this purpose we must:

1) *work out the derivative* $f'(x)$;

2) *find the values of* x *in the interval* (a, b) *for which* $f'(x)$ *vanishes. Let these values be*

$$c_1, c_2, \ldots, c_k;$$

3) *find the sign of the derivative in each of the intervals* (a, c_1), (c_1, c_2), $\ldots$, (c_k, b). *This will automatically solve the question as to whether or not the derivative changes sign on passing (from left to right) through each of the points* $c_1, c_2, \ldots, c_k$; *and if it does change sign, we shall know whether from* $+$ *to* $-$ *or from* $-$ *to* $+$. *A change in the sign of the derivative from* $+$ *to* $-$ *indicates that the function has a maximum at the point in question. A change in the sign of the derivative from* $-$ *to* $+$ *indicates that the function has a minimum for the value of* x *concerned. If the sign of the derivative remains unchanged, the function has no extremum at the point.*

It may be noted here that, since the derivative does not change sign in each portion of the interval, we can use whatever value of x we like from a given portion in order to find the sign of the derivative within that portion.

We shall explain the above with some examples.

Example 1. Find the values of the argument for which the function

$$y=2x^3-3x^2-12x+21$$

has extrema.

Solution. 1) We find the derivative is given by

$$y'=6x^2-6x-12=6\,(x^2-x-2)\,.$$

2) We find the values of x for which the derivative vanishes by solving the equation

$$x^2-x-2=0\,.$$

The roots of this equation are -1 and 2.

To investigate the sign of the derivative, the expression for it may usefully be factorized to give

$$y'=6(x+1)\,(x-2)\,.$$

3) The function y is defined in the interval $(-\infty,\ +\infty)$. This total interval is split by the values -1 and 2 of the argument for which the derivative vanishes into three sub-intervals $(-\infty,\ -1)$, $(-1\,,2)$, $(2\,,+\infty)$.

By finding the signs of the separate factors $(x+1)$ and $(x-2)$ for $x=-2$, $x=0$, $x=3$, we obtain the sign of the derivative in each of our three sub-intervals:

in the interval $(-\infty\,,\ -1)$ the sign of $y'=(-)\cdot(-)=+$,

,, ,, ,, $(-1\,,\ 2)$,, ,, $=(+)\cdot(-)=-$,

,, ,, ,, $(2\,,\ +\infty)$,, ,, $=(+)\cdot(+)=+$.

We see from this that the derivative changes its sign from $+$ to minus on passing through $x=-1$, and from minus to $+$ on passing through $x=2$. Hence $x=-1$ corresponds to a maximum of the function, and $x=2$ to a minimum.

Example 2. Investigate the maxima and minima of the function

$$y=x^3.$$

Solution. 1)

$$y'=3x^2,$$

2) $3x^2=0$, whence $x=0$.

3) Since the derivative remains positive for $x<0$ and $x>0$, it does not in fact change sign on passing through $x=0$. Hence the function does not possess an extremum at $x=0$. The result is shown in a visual form by the graph of the function (fig. 60).

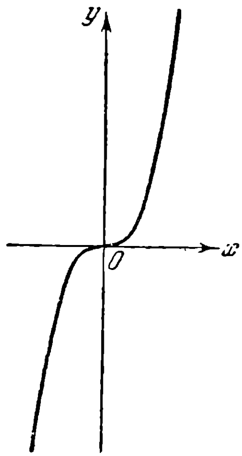


FIG. 60

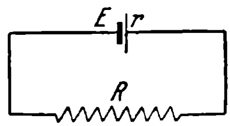


FIG. 61

Example 3. The current I in the circuit of fig. 61 is given by Ohm's law as

$$I=\frac{E}{R+r},$$

where R is the external and r the internal resistance. The power developed in the load R is given by the familiar expression

$$P=I^2 R=\frac{E^2 R}{(R+r)^2}.$$

We shall find the value of R for which the power is a maximum.

Solution. We have to investigate the function P of the independent variable R for maxima and minima.

Using our rule, we have the following steps:

1) $$P'=E^2\frac{(R+r)^2-2(R+r)R}{(R+r)^4}=E^2\cdot\frac{r-R}{(R+r)^3}.$$

2) $r-R=0$, whence $R=r$.

3) The derivative $P'>0$ for $R<r$, whilst $P'<0$ for $R>r$.

Thus the power P is greatest when $R=r$, i.e. when the external and internal resistances of the circuit are equal.

4) The function is not usually expressed directly in problems encountered in practical work. In these cases we have to use the conditions of the problem to set up a relationship between the function and the variables upon which the maxima and minima of the function depend.

The physical nature of the problem often enables us to determine whether a value of the argument for which the derivative vanishes yields a maximum or a minimum. This saves us the necessity of finding the signs of the derivative to the left and right of the value of the argument in question.

Problem 1. We are required to make a lidless box from a square sheet of paper of side a, by cutting squares out of the corners of the sheet and bending the remainder in such a way (fig. 62) that

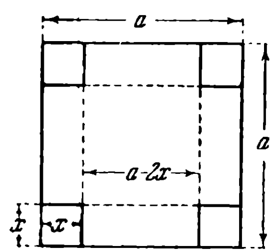


FIG. 62

the box has the greatest possible volume. What must be the length of side of the squares cut out?

Solution. Let the squares cut out have a side of length x. Then x will also be the height of the box. The bottom of the box will be a square of side $a-2x$, so that the volume of the box is $V=(a-2x)^2 x=a^2 x-4ax^2+4x^3$.

We want to find the value of x for which V has a maximum. We thus have to inspect our function V for maxima and minima. By the general rule we have

1) $V' = a^2 - 8ax + 12x^2$;

2) $a^2 - 8ax + 12x^2 = 0$, whence we find $x_1 = \frac{1}{2}a$ and $x_2 = \frac{a}{6}$.

We conclude without further investigation that we obtain a maximum with $x = \frac{a}{6}$, since nothing is left of the sheet if squares of side $\frac{1}{2}a$ are cut out and the volume of the box is zero.

The squares cut out must thus have a side equal to a sixth of the side of the given sheet of paper.

With $x = \frac{a}{6}$, the volume $V = \frac{2a^3}{27}$.

Problem 2. The strength of a rectangular beam is proportional to its width multiplied by the square of its height. Find the dimensions of the strongest beam that can be cut from a cylindrical log of diameter a cm.

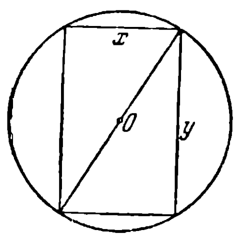


FIG. 63

Solution. Fig. 63 shows a section through the log and beam. Let x be the width and y the height of the beam. Then we have $x^2 + y^2 = a^2$. The strength S of the beam is given by

$$S = kxy^2 = kx(a^2 - x^2) = ka^2x - kx^3,$$

where k is a coefficient of proportionality.

We have thus formed the function that has to be inspected for maxima and minima. We have by the general rule:

1) $S' = k(a^2 - 3x^2)$.

2) $k(a^2 - 3x^2) = 0$, whence $x_1 = \frac{a}{\sqrt{3}}$ and $x_2 = -\frac{a}{\sqrt{3}}$.

We can neglect the negative root, which is meaningless in the context of our problem.

Since the strength of the beam cannot be indefinitely large it must have a maximum; hence the root $x_1 = \frac{a}{\sqrt{3}}$ gives us the value for which S is a maximum; with $x = \frac{a}{\sqrt{3}}$, the height $y = a\sqrt{\frac{2}{3}}$ cm. These are the dimensions of the strongest beam.

Problem 3. Find the height of the cone of maximum volume that can be inscribed in a sphere of radius r.

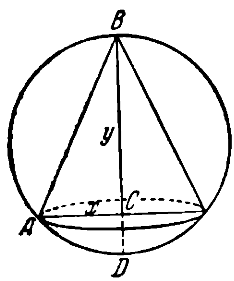


FIG. 64

Solution. Let x be the radius of the base and y the height of the cone (fig. 64). The volume V of the cone will now be

$$V = \frac{1}{3} \pi x^2 y .$$

From fig. 64 we have $x^2 = BC.CD = y(2r - y)$. Hence the volume V is given as follows in terms of the variable y:

$$V = \frac{\pi}{3} y^2 (2r - y) .$$

On investigating this function for maxima and minima, we find that the required height is $y = \frac{4}{3} r$.

5. It must be noted that we have considered the question of extrema only for functions satisfying the following conditions.

1) the function has a continuous derivative in the interval concerned; *

2) the derivative can vanish only at a finite number of points in the interval.

It is precisely with functions of this type that we are mainly concerned in practice.

6. We now show, for the sake of a fuller understanding of the matter, that a function $f(x)$ which is continuous in the interval concerned can also have an extremum at a point where the derivative $f'(x)$ does not exist.

One such case is provided say by the function $f(x)=\sqrt[3]{x^2}$. It may easily be seen that this function has a minimum at $x=0$, although the derivative $f'(x)$ does not exist at this point. For $f(0)=0$, whilst with $x\neq 0$ the quantity $\sqrt[3]{x^2}$ is positive both for $x<0$ and $x>0$.

Hence $f(0)<f(x)$ for $x\neq 0$, which means that the function has a minimum at $x=0$. Whereas, on differentiating $f(x)=x^{2/3}$, we get $f'(x)=\dfrac{2}{3\sqrt[3]{x}}$, and the expression for the derivative becomes meaningless at $x=0$ (division by zero being impossible), so that the rule for differentiating a power function cannot be used for finding the derivative at $x=0$. We must therefore return to the general method of § 42 for finding the required derivative, viz.

1) $f(0)=0$;

2) $f(0+\Delta x)=(0+\Delta x)^{\frac{2}{3}}=(\Delta x)^{\frac{2}{3}}$;

3) $$\begin{array}{r} f(0+\Delta x)=(\Delta x)^{\frac{2}{3}} \\ -\quad f(0)=0 \\ \hline f(0+\Delta x)-f(0)=(\Delta x)^{\frac{2}{3}}; \end{array}$$

* By the theorem proved in § 44, the function itself will be continuous in the interval.

4) $$\frac{f(0+\Delta x)-f(0)}{\Delta x}=\frac{(\Delta x)^{\frac{2}{3}}}{\Delta x}=\frac{1}{\sqrt[3]{\Delta x}};$$

5) $$\lim_{\Delta x\to 0}\frac{f(0+\Delta x)-f(0)}{\Delta x}=\lim_{\Delta x\to 0}\frac{1}{\sqrt[3]{\Delta x}}=\infty .$$

This result shows that the derivative of $f(x)=x^{2/3}$ does not exist at the point $x=0$.

The ratio $\dfrac{f(0+\Delta x)-f(0)}{\Delta x}$ defines the tangent of the angle α of inclination of the secant to the curve $y=x^{2/3}$ through the point $O(0, 0)$, and since $\lim\limits_{\Delta x\to 0}\tan\alpha=\infty$, the limiting position of the secant, i.e. the tangent, forms an angle of $\frac{1}{2}\pi$ with Ox. As shown in fig. 65, the graph of the function has a vertical tangent at the origin.

When deducing the rule for finding the values of x for which the function $f(x)$ has extrema, we assumed that the derivative

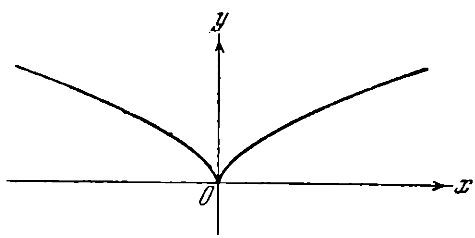


FIG. 65

$f'(x)$ is continuous in the interval (a, b) and can vanish at a finite number of points. It may easily be seen that the arguments of 2. above are completely unchanged in the more general case when (i) the function $f(x)$, continuous in the interval (a, b), has a derivative $f'(x)$ which vanishes or ceases to exist at a finite number of points, and (ii) the derivative is continuous everywhere, apart from the points at which it ceases to exist.

In this more general case, we must find the values of x for which the derivative ceases to exist, as well as those for which it vanishes. We then have to examine each value of x from the point of view of change of sign of the derivative.

The derivative $f'(x) = \frac{2}{3\sqrt[3]{x}}$ of the function $f(x) = \sqrt[3]{x^2}$ changes sign from minus to + on passing through $x=0$. And we have already seen that the function has a minimum at $x=0$.

We now consider $\varphi(x) = \sqrt[3]{x}$. We find its derivative for an arbitrary value of x as $\varphi'(x) = \frac{1}{3\sqrt[3]{x^2}}$. The expression for the derivative is seen to become meaningless at $x=0$. This means that the rule for differentiating a power function cannot be used for finding the derivative at $x=0$. We turn back, as in the previous example, to the general method for evaluating a derivative, and find that

$$\lim_{\Delta x \to 0} \frac{\varphi(0+\Delta x) - \varphi(0)}{\Delta x} = \lim_{\Delta x \to 0} \frac{1}{\sqrt[3]{(\Delta x)^2}} = +\infty ;$$

so that the function has no derivative at $x=0$.

The derivative $\varphi'(x) = \frac{1}{3\sqrt[3]{x^2}} > 0$ both for $x<0$ and $x>0$, so that there is no change of sign on passing through $x=0$. We conclude from the rule for finding extrema that the function has neither a maximum nor a minimum at $x=0$. This is supported by direct inspection of $\varphi(x)$ itself; for $\varphi(0)=0$ is neither a maximum nor a minimum as compared with neighbouring values, since $\varphi(x)-\varphi(0)=\sqrt[3]{x}$ is negative for $x<0$ and positive for $x>0$.

We conclude from the fact that $\lim\limits_{\Delta x \to 0} \frac{\varphi(0+\Delta x)-\varphi(0)}{\Delta x} = +\infty$ that the graph of $y=\sqrt[3]{x}$ has a vertical tangent at the origin (fig. 66).

Note. We proved in § 44 that a function can be continuous at a given point without having a derivative at that point. The functions just considered provide examples of this fact.

§ 63. **Derivative of the second order. Significance of second derivative in Mechanics.** 1. When *the derivative of a function vanishes at some point x,* the existence of an extremum and its actual type (maximum or minimum) can be established *in certain cases* by a different method from that of investigating the change of sign of the derivative on passing through the point. This second method requires the introduction of the new concept of derivative of the second order.

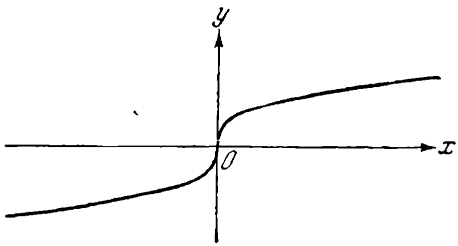


FIG. 66

Let the function $y=f(x)$ have the derived function $y'=f'(x)$. This derived function can in turn have a derivative with respect to its argument x. The derivative of the first derivative $y'=f'(x)$ is known as the *derivative of the second order* or simply as the *second derivative* of the given function $f(x)$, and is written symbolically as

$$y'' \text{ or } f''(x)\,.$$

The second derivative may itself have a derivative. This would be the *derivative of the third order,* or more briefly, the *third derivative* of the given function $y=f(x)$, written $y'''=f'''(x)$, and so on.

Example. Find the second derivative of the function $y=e^{ax}$.

Solution.

$$y'=ae^{ax}\,; \qquad y''=a^2 e^{ax}\,.$$

2. Let a particle move in a straight line in accordance with the law expressed by $s=f(t)$. As we already know (§ 40), the velocity v of the particle at the instant t is defined by the derivative of the path s with respect to time t, i.e.

$$v=s'=f'(t)\,.$$

In mechanics, the *average acceleration* (j_{av}) of the particle in the interval of time Δt (for a given value of t) is defined as the ratio of the increment Δv of the velocity v to the corresponding time interval Δt, i.e.

$$j_{av}=\frac{\Delta v}{\Delta t}.$$

The limit of this ratio as $\Delta t \to 0$ gives the acceleration j at the instant t, i.e.

$$=\lim_{\Delta t \to 0} \frac{\Delta v}{\Delta t}.$$

By definition of derivative, $\lim\limits_{\Delta t \to 0} \frac{\Delta v}{\Delta t}=v'$. And since $v=s'=f'(t)$, we have $v'=s''=f''(t)$. Hence

$$j=s''=f''(t),$$

i.e. *the acceleration of a moving particle is given by the second derivative of the path s with respect to time t.*

Example. A particle moving along a straight line covers a distance of s metres after t seconds in accordance with the formula

$$s=4t^3+2t^2+3t.$$

Find the acceleration: (a) at the initial instant $t=0$; (b) after five seconds ($t=5$).

Solution. The acceleration is the second derivative of the distance s with respect to time t. On twice differentiating the function s with respect to t, we obtain

$$s'=12t^2+4t+3,$$
$$s''=j=24t+4.$$

We have found the acceleration at any instant t. On now substituting $t=0$ and $t=5$ in the expression obtained, we find

(a) $j_{t=0}=4$ m/sec^2, (b) $j_{t=5}=124$ m/sec^2.

§ 64. Second rule for finding the extrema of a function. Let the first derivative of a given function $f(x)$ vanish at $x=c$,

i.e. $f'(c)=0$. Furthermore, let $f(x)$ have a continuous and positive second derivative at the point c, i.e. $f''(c)>0$.

Since $f''(x)$ is continuous at $x=c$, a slight change in the argument x from $x=c$ will produce only a slight change in the value of the second derivative. Hence $f''(x)$ will retain the sign that it has at $x=c$ in some neighbourhood of this point, i.e. there will be a certain interval in which $f''(x)$ is positive.

We know that if the derivative of a function is positive in an interval, the function is increasing in this interval (§ 61). Now $f''(x)$ is the derivative of the first derivative $f'(x)$. Hence $f'(x)$ is increasing from the left to the right of $x=c$ in some interval about $x=c$. Also, since $f'(c)=0$, $f'(x)$ must be negative in some interval to the left of $x=c$ and positive in some interval to the right.

It follows that, if the second derivative $f''(c)>0$ (and is continuous at the point c), the first derivative $f'(x)$ changes sign from minus to plus on passing through $x=c$. Hence the function $f(x)$ has a minimum at $x=c$ (§ 62, no. 2).

It can be shown in exactly the same way that if $f''(c)<0$, the function has a maximum at $x=c$.

This method for finding the extremals of a function cannot be used if $f''(x)$ also vanishes at the point where $f'(x)=0$. In this case we have to solve the problem by returning to the basic rule outlined in § 62.

The results obtained can be set out as follows:

x	$f'(x)$	$f''(x)$	$f(x)$
		$-$	maximum
c	0	$+$	minimum
		0	rule unusable

Example 1. Investigate the maxima and minima of the function

$$y=x^4-2x^2+2\,.$$

Solution. 1) We find the first derivative as

$$y'=4x^3-4x=4x(x^2-1)$$

or

$$y' = 4x(x+1)(x-1).$$

2) We find the roots of the first derivative, i.e. the values of x for which y' vanishes:

$$x=0, \quad x=-1, \quad x=+1.$$

3) We find the second derivative as

$$y'' = 12x^2 - 4.$$

4) On substituting the roots of the first derivative in the expression for the second derivative, we get

$$\begin{array}{ll} \text{for } x=0 & y''<0 \\ \text{for } x=-1 & y''>0 \\ \text{for } x=+1 & y''>0. \end{array}$$

Thus the function $y=x^4-2x^2+2$ has a maximum at $x=0$ and minima at $x=-1$ and $x=+1$.

Example 2. Investigate the maxima and minima of the function $y=x^4$.

Solution. 1) $y'=4x^3$; 2) $4x^3=0$, whence we find the root $x=0$e 3) $y''=12x^2$; 4) $y''=0$ at $x=0$.

We therefore turn back to the first method to find that

$$\begin{array}{ll} \text{at } x<0 & y'<0, \\ \text{at } x>0 & y'>0. \end{array}$$

The function $y=x^4$ therefore has a minimum at $x=0$. Th. graph of the function (a parabola of the fourth order) is shown in fig. 67.

§ **65. Concavity of a curve at a point.** 1. The graphs of certain functions will have been traced when learning about functions in the course of elementary algebra. The graphs are obtained in this case by plotting separate points, the choice of which is strictly speaking arbitrary, rather than based on any definite arguments. With this method, certain important features of the curve, i.e. of the behaviour of the function, can escape our attention, due to

occurring so to speak "between" the points that we have decided to plot.

The methods of the differential calculus provide us with advance information regarding peculiarities in the behaviour of a given function: we can find the intervals in which it is increasing or decreasing, and the points at which it has extrema. As a result,

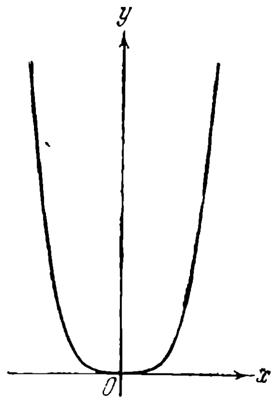


FIG. 67

a graph can be drawn that gives a far more accurate geometrical picture of the variation of the function. A still more detailed picture can be obtained, however, if we can discover in which direction the various portions of the graph are concave, and at which points the concavity changes. These questions will be discussed in the present article and the next.

2. *A curve is said to be concave upwards or convex at a point M if every point of some arc of the curve (however small) to the left and right of M lies above the tangent at M.*

Similarly, *a curve is said to be concave downwards at a point M if every point of some arc of the curve (however small) to the left and right of M lies below the tangent at M.*

Figures 68 and 69 give a visual illustration of these definitions: fig. 68 shows the case of a curve concave upwards at the point M, and fig. 69 the concave downwards case.

We establish an analytic test for finding whether a curve is concave upwards or concave downwards at a given point.

Let the curve be given by the equation $y=f(x)$. Let c be the abscissa of the point M, and let the function $f(x)$ have a second derivative $f''(c)$ at $x=c$.

Our test is expressed as the following theorem:

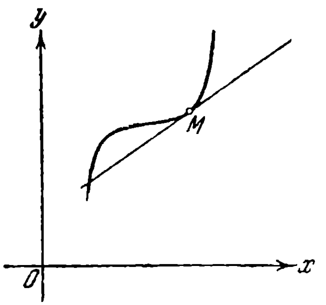


FIG. 68

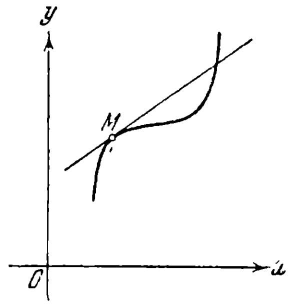


FIG. 69

THEOREM (sufficient tests to determine the concavity of a curve at a point). *Let the curve be the graph of the function $y=f(x)$, and let c be the abscissa of the point M of the curve. If $f''(c)>0$,*

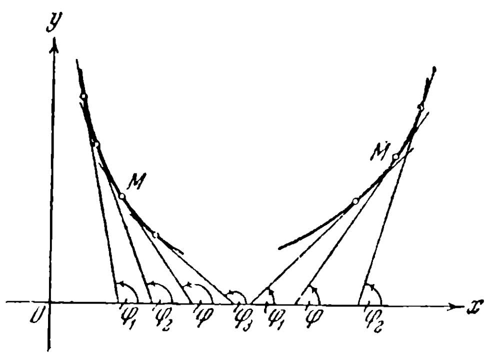


FIG. 70

the curve is concave upwards at the point M, whilst if $f''(c)<0$, it is concave downwards.

We shall take the case when the second derivative of our function $f(x)$ is continuous at the point c.

Suppose $f''(c)>0$. Since $f''(x)$ is continuous at $x=c$, a small variation in x leads to a change in the second derivative which is

likewise small. Hence $f''(x)$ remains positive in a certain neighbourhood of the point c (to the left and right of c). It follows by the test for an increasing function (§ 61) that the first derivative $f'(x)$ is increasing in this interval.

The first derivative is the slope of the tangent to the curve at the given point, i.e. the tangent of the angle φ which the tangent forms with the positive direction of Ox. As $\tan\varphi$ increases, the angle φ increases (anti-clockwise). Hence the angle of inclination of the tangent to the curve increases as its point of contact approaches M from the left, and continues to increase as the point of contact moves away from M to the right (however small the portion of the curve concerned). At the same time, we can see from fig. 70 that an arc of the curve representing $y=f(x)$ must, at least as regards

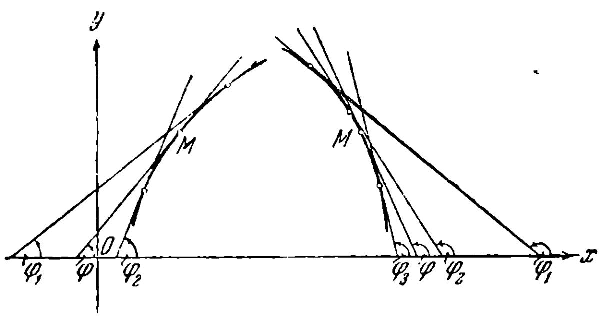


FIG. 71

some small portion close to M, be situated above the tangent, i.e. the curve must be concave upwards at the point M.*

It can be seen from precisely similar arguments that, if $f''(c)<0$, the curve representing $y=f(x)$ must be concave downwards at the point M with abscissa c (fig. 71).

If the curve of $y=f(x)$ is concave upwards (concave downwards) for all values of x in the interval (a, b), the curve is said to be concave upwards (concave downwards) in the interval (a, b).

---

* To understand this idea, the student is recommended to reproduce fig.70 by first drawing a series of straight lines forming angles $\varphi_1<\varphi_2<\varphi_3\ldots$ with the positive direction of Ox. It will then be clear that any curve touching these tangents must lie above them.

§ 66. Points of inflexion. 1. It may happen that the graph of $y=f(x)$ is convex downwards in certain intervals and concave in others. For instance, the curve shown in fig. 72 is concave upwards in the intervals (a, b) and (c, d), and concave downwards in the interval (b, c).

A point M of a curve is called a point of inflexion if it separates a concave upwards portion from a concave downwards portion of the curve.

In fig. 72, B and C are points of inflexion. It is clear that a curve having a tangent at a point of inflexion must pass from one side of the tangent to the other at this point.

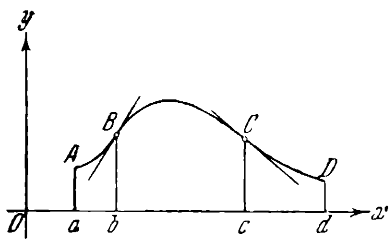


FIG. 72

The points of inflexion of a function form, along with the extrema, the base points for drawing a graph of the function. We must therefore discover how to recognize the points of inflexion of a curve from its equation $y=f(x)$.

2. We suppose that the function $y=f(x)$ is continuous in the interval (a, b) and has continuous derivatives $f'(x)$ and $f''(x)$ in this interval.

Let us suppose that the second derivative $f''(x)$ does not vanish for any x in (a, b). Since it is continuous, it must retain the same sign throughout (a, b) (see § 62, no. 2, where it is shown that, if a continuous function changes sign in an interval, it must vanish for some value of the argument in that interval). Hence the curve representing the given function must be either concave upwards everywhere in the interval (a, b) or concave downwards everywhere in (a, b), so that there can be no points of inflexion.

It follows that, if $y=f(x)$ has a graph with points of inflexion in (a, b), it can only be for those values of x for which $f''(x)$ vanishes.

Now let $f''(x)$ vanish in (a, b) but only at a finite number of points, say $c_1 < c_2 < \ldots . < c_k$. Now $f''(x)$ has a fixed sign in each of the intervals (a, c_1), (c_1, c_2), ..., (c_k, b). We take any one of the points $c_1, c_2, \ldots, c_k$, say c_1. Suppose $f''(x)>0$ in (a, c_1), and $f''(x)<0$ in (c_1, c_2). Then the graph of $y=f(x)$ is concave upwards in (a, c_1) and concave downwards in (c_1, c_2) and therefore has a point of inflection at $x=c_1$. In the same way, we find that the curve has a point of inflexion at $x=c_1$ if we suppose that $f''(x)<0$ in (a, c_1) and $f''(x)>0$ in (c_1, c_2).

Finally, let the sign of $f''(x)$ be the same in both (a, c_1) and (c_1, c_2). Then the graph of the function is either concave upwards or concave downwards in both intervals, i.e. the curvature is in the same direction throughout the interval (a, c_2), so that there can be no point of inflexion at $x=c_1$.

The same arguments allow us to decide the question of points of inflexion on the graph of the function for the remaining values $c_2, c_3, \ldots, c_k$ of the argument.

Our discussion leads to the following rule for finding the values of x at which the graph of $y=f(x)$ has points of inflexion. We must:

1) *evaluate the second derivative $f''(x)$ of the given function $f(x)$, the graph of which is under investigation*;

2) *find the values of x for which $f''(x)$ vanishes in the interval (a, b); let these values be $c_1, c_2, \ldots, c_k$*;

3) *find the sign of $f''(x)$ in each of the intervals (a, c_1), (c_1, c_2), ..., (c_k, b), for which purpose it is sufficient to find the sign of $f''(x)$ for any one value of x from each interval* (see § 62, no. 3). *We shall now have decided simultaneously the question as to whether or not the sign of $f''(x)$ changes on passing through each of the points $c_1, c_2, \ldots, c_k$. A change in the sign of $f''(x)$ shows that the graph of $f(x)$ has a point of inflexion for the value of x in question. If the sign of $f''(x)$ does not change, there is no point of inflexion.*

To put the matter briefly, we carry out the same investigation with regard to the second derivative to find the points of inflexion

$f'(x)$ multiplied by the differential dx of the argument, i.e.

$$dy = f'(x)\,dx.$$

Finding the differential thus amounts to finding the derivative of the given function and multiplying by dx.

There can be no difficulty about this, since we already know how to find derivatives.

The basic formulae for differentiation given in § 45 can be written in terms of differentials as follows: if u and v are functions of x having derivatives for the value of x in question, we have *

$$\text{I.}\quad d(u \pm v) = (u' \pm v')\,dx = u'dx \pm v'dx = du \pm dv$$

(since $u'dx$, i.e. the product of dx and the derivative of u, is the differential of u, i.e. du; similarly, $v'dx = dv$.)

$$\text{II.}\quad d\,(uv) = (u'v + v'u)\,dx = v\,(u'\,dx) + u\,(v'\,dx)$$
$$= v\,du + u\,dv.$$

$$\text{III.}\quad d\,(cu) = cu'\,dx = c\,du.$$

$$\text{IV.}\quad d\left(\frac{u}{c}\right) = \frac{1}{c}\,u'\,dx = \frac{1}{c}\,du = \frac{du}{c}.$$

$$\text{V.}\quad d\left(\frac{u}{v}\right) = \frac{u'v - v'u}{v^2}\,dx = \frac{v\,(u'\,dx) - u\,(v'\,dx}{v^2}$$
$$= \frac{v\,du - u\,dv}{v^2}.$$

$$\text{VII.}\quad dc = 0 \quad (c \text{ is a constant}).$$

$$\text{IX.}\quad d\,(x^a) = ax^{a-1}\,dx \quad (a \text{ is a constant}).$$

$$\text{X.}\quad d\,(\sin x) = \cos x\,dx\,.$$

$$\text{XI.}\quad d\,(\cos x) = -\sin x\,dx\,.$$

$$\text{XII.}\quad d\,(\tan x) = \frac{dx}{\cos^2 x}.$$

---

* Some of the formulae of § 45 are omitted here.

XIV. $d\,(\log_a x)=\frac{1}{x}\ \mathrm{og}_a\, e\,dx.$

XV. $d\,(\ln x)=\frac{dx}{x}.$

XVI. $d\,(a^x)=a^x \ln a\,dx.$

XVII. $d\,(e^x)=e^x\,dx.$

XVIII. $d\,(\text{arc}\sin x)=\frac{dx}{\sqrt{1-x^2}}.$

XIX. $d\,(\text{arc}\cos x)=-\frac{dx}{\sqrt{1-x^2}}.$

XX. $d\,(\text{arc}\tan x)=\frac{dx}{1+x^2}.$

XXI. $d\,(\text{arc}\cot x)=-\frac{dx}{1+x^2}.$

Example. Find the differential of the function

$$y=\tan(1+x^2).$$

Solution. Since

$$y'=\frac{1}{\cos^2(1+x^2)}\cdot 2x,$$

we have

$$dy=\frac{2x}{\cos^2(1+x^2)}\,dx.$$

The term differentiation is applied equally to the process of finding the differential of a given function $y=f(x)$ and the process of finding the derivative.

§ 72. The application of differentials to approximate evaluations. 1. The differential $dy=f'(x)\,\Delta x$ of the function $y=f(x)$ is *linearly* dependent on Δx. The increment, on the other hand, is defined by the relationship

$$\Delta y = dy + \alpha \Delta x$$

(where $\alpha \to 0$ as $\Delta x \to 0$), which in general represents a more complicated function of Δx. Due to this, evaluation of the increment of a function is far more difficult than evaluation of its differential.

For values of Δx close to zero, the values of α are also close to zero. Hence substitution of dy for Δy (for values of Δx close to zero) leads only to very small errors in practice.

The substitution of the differential dy for the increment Δy in approximate evaluations is based on the facts, (i) that it is easy to work out the differential; (ii) that the difference between Δy and dy is insignificant (for values of Δx close to zero).

2. This substitution is extremely useful for working out errors.

Suppose, for instance, that the quantity x is found by direct measurement, whilst y is obtained in accordance with the formula $y = f(x)$. An error Δx always creeps in when measuring x, which leads in turn to an error Δy in working out y. Since the error Δx will always be extremely small with accurate measurements, we can write

$$\Delta y \approx dy = y' \Delta x,$$

i.e. we replace Δy by the differential dy.

The maximum possible error on either side of the indicated value of x can usually be determined from the nature of the measuring equipment. In the case of a slide-rule of length 250 mm, for instance, the error in reading or setting the slider does not exceed 0·1 mm. Thus we generally know the *maximum absolute error* in measuring the quantity x. We denote this error by δx. Then $|\Delta x| \leqq \delta x$. For the maximum absolute error in evaluating y, we now naturally take

$$\delta y = |\, y' \,|\, \delta x.$$

The ratios $\dfrac{\delta x}{|x|}$ and $\dfrac{\delta y}{|y|}$ are termed the *maximum relative errors* in determining x and y respectively. These ratios give a clear indication of the accuracy of the results obtained.

Some examples will show how these ideas are applied.

Example 1. Suppose that the area of a circle is found by first measuring directly its diameter D (with the aid of a micrometer, beam compasses, etc.) then working out the area S from the formula

$$S=\frac{\pi D^2}{4}.$$

Since

$$S'=\frac{\pi}{2}\cdot D,$$

we have

$$\delta S=\frac{\pi}{2}D\delta D$$

and

$$\frac{\delta S}{S}=\frac{\frac{\pi}{2}D\,\delta D}{\frac{\pi}{4}D^2}=2\frac{\delta D}{D}.$$

The maximum relative error in evaluating the area is therefore twice as great as the maximum relative error in measuring the diameter.

Example 2. Find the best angle of deviation φ from the point of view of accuracy of the magnetic needle in a tangent galvanometer when measuring a current.

We know that the current i is proportional to $\tan\varphi$, i.e. $i=k\tan\varphi$, where k is a coefficient of proportionality.

Let the maximum absolute error in reading the angle φ be $\delta\varphi$. Then

$$\delta i=(k\tan\varphi)'\,\delta\varphi=\frac{k\delta\varphi}{\cos^2\varphi}$$

and

$$\frac{\delta i}{i}=\frac{k\,\delta\varphi}{\cos^2\varphi}:(k\tan\varphi)=\frac{\delta\varphi}{\sin\varphi\cdot\cos\varphi}=\frac{2\delta\varphi}{\sin 2\varphi}.$$

We see that the accuracy will be greatest when $\sin 2\varphi=1$, i.e. $\varphi=\frac{\pi}{4}$, and least for values of $\sin 2\varphi$ close to zero, i.e. for values

of φ near to 0 and $\frac{1}{2}\pi$, in which case the relative error increases indefinitely.

One must therefore avoid taking readings at angles of deviation close to 0 and $\frac{1}{2}\pi$.

Suppose $\delta\varphi=0{\cdot}5°$. On expressing the angle in radian measure, we find

$$\delta\varphi=0{\cdot}5\times\frac{\pi}{180}=\frac{1}{2}\times 0{\cdot}01745$$

and

$$\frac{\delta i}{i}=\frac{2\times\frac{1}{2}\times 0{\cdot}01745}{\sin 2\varphi}=\frac{0{\cdot}01745}{\sin 2\varphi}.$$

Say we take a reading at 30°, then $\sin 2\varphi=\sin 60°=\frac{1}{2}\sqrt{3}$ $=\frac{1}{2}=\cdot 1{\cdot}73205$ and

$$\frac{\delta i}{i}=\frac{2\times 0{\cdot}01745}{1{\cdot}73205}$$

i.e. about 2%.

3. Substitution of the differential dy of a function $y=f(x)$ for the increment Δy is also used for obtaining approximation formulae which are often employed in practical work.

For instance, the following approximation formula is often used with small values of $|h|$:

$$(1+h)^a\approx 1+ah\ .$$

The formula is deduced in the following manner.

We take the function $y=x^a$. We set $x=1$, $x=1+h$ and form the increment of the function $\Delta y=(1+h)^a-1$. We now replace the increment by the differential $dy=ax^{a-1}dx$. With $x=1$ and $dx=h$. we get approximately

$$(1+h)^a-1\approx ah\ ,$$

whence we find

$$(1+h)^a\approx 1+ah\ .$$

Similarly, for small values of φ we can write

$$\sin \varphi \approx \varphi .$$

For, on taking the function $y=\sin x$ and setting $x=0$, $x=\phi$, we obtain the increment Δy as

$$\Delta y=\sin \varphi-\sin 0=\sin \varphi .$$

On replacing Δy by the differential $dy=\cos x\, dx$, with $x=0$ and $dx=\phi$ we arrive at the approximation

$$\sin \varphi \approx \varphi.$$

An approximation can be similarly deduced for $\ln (1+h)$ with small values of h.

If $y=\ln x$, $dy=\frac{dx}{x}$. By setting $x=1$, $x=1+h$, we form the increment Δy as

$$\Delta y=\ln (1+h)-\ln 1=\ln (1+h) ,$$

whence with $x=1$ and $dx=h$ we obtain approximately

$$\Delta y=\ln (1+h) \approx \frac{h}{1}=h .$$

§ 73. **Differential of arc length.** Let a curve be given by the equation $y=f(x)$, and let s denote the variable length of arc measured from a fixed point A to a point M which can move along the curve (fig. 81).

Then s will be a function of the abscissa x of the point M, since to every value of x there corresponds a definite position of M on the curve, and consequently a definite length s of the arc AM. Let us find the differential of this function s.

We know that the differential of a function is the product of the derivative and the differential of the argument (§ 69). Hence we can write

$$ds=s'\, dx.$$

Thus our problem amounts to finding the derivative s' of the arc.

We do this by giving a chosen value x of the argument the increment $\Delta x = dx$, as a result of which the point M moves to the position M' and the arc $s = \overset{\frown}{AM}$ receives the increment $\Delta s = \overset{\frown}{MM'}$. The derivative of the arc is the limit of the ratio $\frac{\Delta s}{\Delta x}$ as $\Delta x \to 0$. To find this limit, we transform the ratio by multiplying and dividing it by the length of the chord $\overline{MM'}$. Thus

$$\frac{\Delta s}{\Delta x} = \frac{\Delta s}{\overline{MM'}} \cdot \frac{\overline{MM'}}{\Delta x}.$$

We now pass to the limit:

$$\lim_{\Delta x \to 0} \frac{\Delta s}{\Delta x} = \lim_{\Delta x \to 0} \frac{\Delta s}{\overline{MM'}} \cdot \lim_{\Delta x \to 0} \frac{\overline{MM'}}{\Delta x}.$$

We shall accept without proof that $\lim\limits_{\Delta x \to 0} \frac{\Delta s}{\overline{MM'}} = 1$. A rigorous proof of this is rather difficult for an elementary course, although

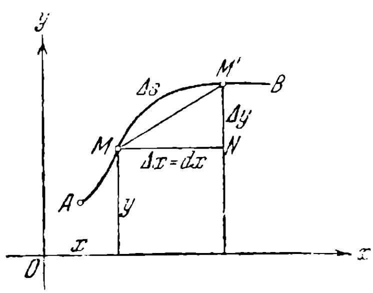


FIG. 81

the result is sufficiently clear, since when M' approaches M along the curve, the increment of arc Δs resembles more and more closely the chord $\overline{MM'}$.

The result was established perfectly rigorously for the case of a circle in § 53, where it was shown that $\lim\limits_{x \to 0} \frac{\sin x}{x} = 1$.

Since $\overline{MM'}=\sqrt{(\Delta x)^2+(\Delta y)^2}$, we obtain

$$\lim_{\Delta x\to 0}\frac{\Delta s}{\Delta x}=\lim_{\Delta x\to 0}\frac{\overline{MM'}}{\Delta x}=\lim_{\Delta x\to 0}\frac{\sqrt{(\Delta x)^2+(\Delta y)^2}}{\Delta x}$$

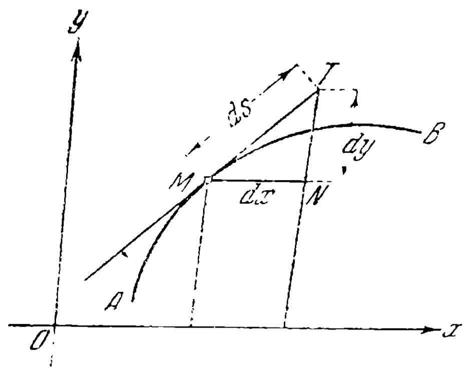


FIG. 82

or
$$\lim_{\Delta x\to 0}\frac{\Delta s}{\Delta x}=\lim_{\Delta x\to 0}\sqrt{1+\left(\frac{\Delta y}{\Delta x}\right)^2}.$$

In view of the fact that $\lim\limits_{\Delta x\to 0}\left(\dfrac{\Delta y}{\Delta x}\right)=y'$, we have $s'=\sqrt{1+y'^2}$, so that

$$ds=\sqrt{1+y'^2}\,dx. \tag{6}$$

We can write this in another way by taking dx inside the square root sign, i.e.

$$ds=\sqrt{dx^2+dy^2}. \tag{6*}$$

The geometrical meaning of the differential of an arc may easily be seen. As we know, the differential dy of the ordinate of the curve $y=f(x)$ is equal to the increment of the ordinate of the tangent corresponding to the increment $\Delta x=dx$ (§ 70). And as can readily be seen from fig. 82,

$$ds=\sqrt{dx^2+dy^2}=\sqrt{\overline{MN}^2+\overline{NT}^2}=\overline{MT},$$

i.e. the differential of the arc length is equal to the length of the

tangent from its point of contact to its point of intersection with the ordinate of the point whose abscissa is $x+dx$.

§ 74. **Curvature.** One of the characteristic elements of a curve is the degree to which it bends or curves round at different points.

As we pass from one point M to another point M' on a curve, the position of the tangent to the curve varies simultaneously, i.e. the tangent rotates. We are speaking here of curves in the ordinary non-mathematical sense, as distinct from straight lines. The tangent to a straight line coincides with the line itself at every point and its direction is therefore always the same. It is evident that, the greater the angle through which the tangent rotates on passing from the point M to M', the more the arc of the curve will differ from a straight line, and the greater will be its curvature.

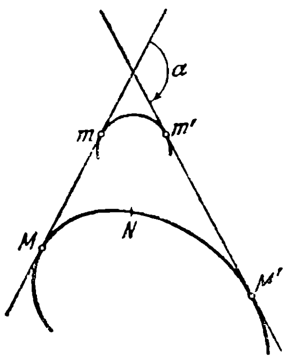


FIG. 83

The curvature of an arc cannot be completely defined, however, simply by the angle through which the tangent rotates on passing from one end of the arc to the other. Figure 83 illustrates two arcs $\overset{\frown}{mm'}$ and $\overset{\frown}{MM'}$, which clearly have different curvatures although the angle of rotation of the tangent is the same. To establish a measure for the curvature, we must also take into account the length of the arc. As is clear from fig. 83, the shorter the length of arc for a given angle of rotation of the tangent, the greater the curvature. Hence it seems natural to take for the *average curvature* of an arc

the ratio of the angle between the tangents at the ends to the length of the arc, or in other words, the angle of rotation of the tangent per unit length of arc. The mean curvature of the arc $\overset{\frown}{MM'} = \dfrac{\alpha}{\overset{\frown}{MM'}}$.

The angle α is assumed always to be expressed in radians.

We shall work out as an example the mean curvature of a circle of radius R. For any length of arc of the circle, the angle α between the tangents at the ends is equal to the angle between the radii drawn to the ends (fig. 84). The length of arc $\overset{\frown}{AB} = \alpha R$. Thus the

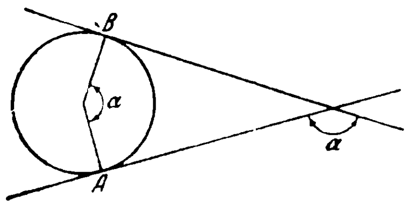


FIG. 84

average curvature is the same for any portion of the circle and is equal to $\dfrac{\alpha}{\alpha R} = \dfrac{1}{R}$, i.e. the mean curvature of a circle is the reciprocal of its radius. Obviously, the smaller the radius, the greater the mean curvature.

The mean curvature is characteristic of the curvature of the arc as a whole. The curvature may be different, however, over different portions of the same arc. For instance, it is clear to the eye that the part $\overset{\frown}{MN}$ of arc $\overset{\frown}{MM'}$ in fig. 83 has a greater curvature than the part NM'. All this suggests that the mean curvature will give a better indication of the extent to which an arc curves at different individual points, the smaller the length of arc in question. This leads us naturally to taking *the curvature K of a curve at a point M* as the limit of the mean curvature of the arc $\overset{\frown}{MM'}$ as $\overset{\frown}{MM'}$ tends to zero, i.e.

$$K = \lim_{\overset{\frown}{MM'} \to 0} \frac{\alpha}{\overset{\frown}{MM'}}.$$

The reader will observe that the passage from the mean curvature to the curvature at a point is precisely analogous to the passage from the average velocity of a particle to the velocity at a given instant.

We now deduce the formula for the curvature K at a point. The curve will usually be given by its equation $y=f(x)$. Since we

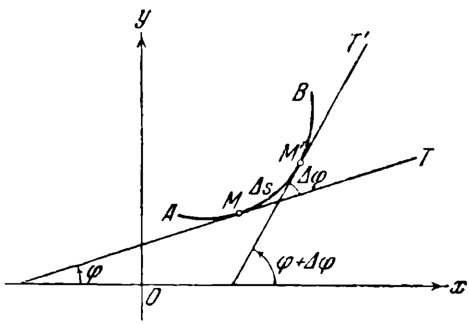


FIG. 85

want to find the curvature at a point of the curve, we must assume that the co-ordinates of the point are given in addition to the equation $y=f(x)$. To deduce the formula for the curvature now implies expressing $\lim \frac{\alpha}{\overset{\frown}{MM'}}$ in terms of the co-ordinates x, y of the point M whilst bearing in mind that $y=f(x)$.

Given the curve $y=f(x)$, let the co-ordinates of the point M at which we require to find the curvature be x, y. Let x take the increment Δx. For the abscissa $x+\Delta x$, the corresponding point of the curve will be M'. We draw the tangents MT and $M'T'$ at M and M' (fig. 85). If φ denotes the angle of inclination of the tangent at M, this angle will receive the increment $\Delta\varphi$ on passing from M to M'. The increment $\Delta\varphi$ is also the angle through which the tangent rotates. Let arc $\overset{\frown}{AM}=s$ and arc $\overset{\frown}{MM'}=\Delta s$. Then we have

$$K=\lim_{\Delta x\to 0}\frac{\Delta\varphi}{\Delta s}\,. \qquad (6^*/3)$$

We divide top and bottom of the fraction by Δx and pass to the limit as $\Delta x \to 0$ and obtain

$$K = \lim_{\Delta x \to 0} \frac{\dfrac{\Delta \varphi}{\Delta x}}{\dfrac{\Delta s}{\Delta x}} = \frac{\lim\limits_{\Delta x \to 0} \dfrac{\Delta \varphi}{\Delta x}}{\lim\limits_{\Delta x \to 0} \dfrac{\Delta s}{\Delta x}} = \frac{\varphi'}{s'} .$$

We found above (§ 73) that $s' = \sqrt{1 + y'^2}$. The derivative ϕ' may be worked out in the following manner:

We know that $\tan \varphi = y'$ (§ 43). Hence $\varphi = \arctan y'$. Since y' is a function of x, $\arctan y'$ must be a function of a function of x, and differentiation with respect to x gives us

$$\varphi' = \frac{y''}{1 + y'^2} .$$

On substituting for s' and φ' in the expression for the curvature, we obtain

$$K = \frac{y''}{(1 + y'^2)^{3/2}} . \tag{7}$$

We have thus achieved our purpose of finding an expression for K in terms of the given function $y = f(x)$ and the given value of x. Since the square root in the denominator is always taken with the plus sign in front, the curvature may be positive or negative depending on the sign of y''. This shows that the curvature is positive at points where the curve is concave upwards, and negative where the curve is concave downwards.

§ 75. **Circle of curvature and radius of curvature.** Let the curvature of a given curve AB be equal to K at the point M (fig. 86). We draw the tangent and normal to the curve at M.* If we now draw circles through M with centres on the normal in the direction in which the curve is concave, they will all have a tangent in common

* The normal to the curve at M is the straight line through M perpendicular to the tangent at M.

with the curve AB at M and will all be concave in the same direction as the curve. Among these circles there will be one that has the same curvature K as the curve at the point M. It follows from the

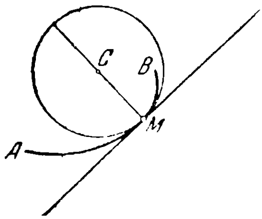


FIG. 86

connexion between the curvature and the radius of a circle (§ 74) that the circle in question must have a radius equal to the absolute value of $\frac{1}{K}$.

The circle obtained in this way is known as the *circle of curvature* of the curve at the point M. The reciprocal of the curvature at M, i.e. $R=\frac{1}{K}$, is termed the *radius of curvature*, whilst the centre C of the circle of curvature is termed the *centre of curvature* of the curve at the point M.

Just as the tangent indicates the rise of a curve at a given point, the circle of curvature gives a visual picture of the curvature at a given point. The circle of curvature touches the curve more closely at M than any other circle and gives a good approximation to the curve over a small arc about M.

The expression for the radius of curvature follows very simply from the equation

$$R=\frac{1}{K};$$

since

$$K=\frac{y''}{(1+y'^2)^{3/2}}\,.$$

we have

$$R=\frac{(1+y'^2)^{3/2}}{y''}\,. \tag{8}$$

§ 76. Examples of evaluating radii of curvature. 1. Find the radius of curvature of the rectangular hyperbola $xy=12$ at the point $(3, 4)$.

Solution. We have $y=\frac{12}{x}$, $y'=-\frac{12}{x^2}$, $y''=\frac{24}{x^3}$. With $x=3$, the first and second derivatives have the values

$$y'=-\frac{12}{9}=-\frac{4}{3}, \qquad y''=\frac{24}{27}=\frac{8}{9},$$

hence

$$R=\frac{\left(1+\frac{16}{9}\right)^{3/2}}{\frac{8}{9}}=\frac{125}{24}.$$

2. Find the radius of curvature of the ellipse

$$b^2x^2+a^2y^2=a^2b^2$$

at the point $(0, b)$.

Solution. We differentiate with respect to x both sides of the equation of the ellipse, y being regarded as a function of x and y^2 as a function of a function of x. We obtain

$$\frac{2x}{a^2}+\frac{2yy'}{b^2}=0;$$

hence we find for y':

$$y'=-\frac{b^2x}{a^2y}.$$

We differentiate this equation with respect to x, bearing in mind that y is a function of x to give,

$$y''=-\frac{b^2}{a^2}\cdot\frac{y-xy'}{y^2}.$$

Substitution in this of the expression obtained for y' gives us

$$y''=-\frac{b^2}{a^2}\cdot\frac{y+\frac{b^2x^2}{a^2y}}{y^2}=-\frac{b^2}{a^2}\cdot\frac{a^2y^2+b^2x^2}{a^2y^3}=-\frac{b^4}{a^2y^3}$$

(since $a^2y^2+b^2x^2=a^2b^2$). Thus

$$R=\frac{\left(1+\frac{b^4x^2}{a^4y^2}\right)^{3/2}}{-\frac{b^4}{a^2y^3}}=-\frac{(a^4y^2+b^4x^2)^{3/2}}{a^4b^4}.$$

We obtain at the point $(0, b)$, i.e. with $x=0$, $y=b$,

$$R=-\frac{a^2}{b}.$$

3. Find the radius of curvature at any point of the catenary (fig. 87)

$$y=\frac{a}{2}\left(e^{\frac{x}{a}}+e^{-\frac{x}{a}}\right)^*$$

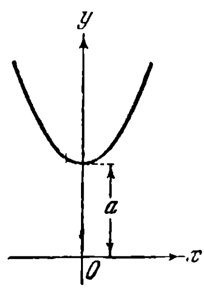


FIG. 87

Solution.

$$y'=\frac{1}{2}\left(e^{\frac{x}{a}}-e^{-\frac{x}{a}}\right) \quad \text{and} \quad \sqrt{1+y'^2}=\frac{1}{2}\left(e^{\frac{x}{a}}+e^{-\frac{x}{a}}\right)=\frac{y}{a}.$$

On differentiating the first derivative we obtain

---

* This is the shape assumed by a flexible non-extensible string suspended from its ends.

$$y'' = \frac{1}{2a}\left(e^{\frac{x}{a}} + e^{-\frac{x}{a}}\right) = \frac{y}{a^2}.$$

Hence

$$R = \frac{\left(\frac{y}{a}\right)^3}{\frac{y}{a^2}} = \frac{y^2}{a}.$$

EXERCISES

On § 71.

Find the differentials of the following functions:

1. $y = 2x^3$. Ans. $dy = 6x^2\,dx$.

2. $y = \sqrt{x}$. Ans. $dy = \frac{dx}{2\sqrt{x}}$.

3. $y = 3x^2 - 6x$. Ans. $dy = 6(x-1)\,dx$.

4. $y = 3x^2 + \frac{1}{x}$. Ans. $dy = \left(6x - \frac{1}{x^2}\right)dx$.

5. $y = (x^3 - a)^2$. Ans. $dy = 6x^2(x^3 - a)\,dx$.

6. $y = ax^3 - bx^2 + cx + d$. Ans. $dy = (3ax^2 - 2bx + c)\,dx$.

7. $y = 2x^{\frac{5}{2}} - 3x^{\frac{2}{3}} + 6x^{-1} + 5$. Ans. $dy = (5x^{\frac{3}{2}} - 2x^{-\frac{1}{3}} - 6x^{-2})\,dx$.

8. $y = (a^2 - x^2)^4$. Ans. $dy = -8x(a^2 - x^2)^3\,dx$.

9. $y = \ln\sqrt{1 - x^3}$. Ans. $dy = \frac{3x^2\,dx}{2(x^3 - 1)}$.

10. $y = (e^x + e^{-x})^2$. Ans. $dy = 2(e^{2x} - e^{-2x})\,dx$.

On § 72.

11. Find the increment and differential of the function $y = 2x^2 - x$ on passing from the value $x = 1$ to $x = 1{\cdot}01$.

Ans. $\Delta y = 0{\cdot}0302$, $dy = 0{\cdot}03$.

12. Find the increment and differential of the function $y=x^3+2x$ on passing from $x=-1$ to $x=-0{\cdot}98$.

Ans. $\Delta y=0{\cdot}098808$, $dy=0{\cdot}1$.

13. Find the increment and differential of the function $y=x^3-$ $-5x^2+80$ on passing from $x=4$ to $x=4{\cdot}001$.

Ans. $\Delta y=0{\cdot}008007001$, $dy=0{\cdot}008$.

14. The period of oscillation of a pendulum is given by

$$T=2\pi\sqrt{\frac{l}{g}},$$

where l is the length of the pendulum and g is the acceleration due to gravity. Find the maximum relative error of the pendulum due to a maximum error of δl in measuring the length l.

Ans. $\dfrac{\delta T}{T}=\dfrac{1}{2}\,\dfrac{\delta l}{l}$.

15. The side of a square, measured to an accuracy of 0·1 m was found to be 4·2 m. What are the maximum absolute and relative errors for the area of the square?

Ans. 0·84 m^2; about 4·8%.

16. The same problem as 15, but for the volume of a cube with a side of 4·2 m.

Ans. 5·292 m^3; about 7·1%.

17. Find the approximate value of $\dfrac{x}{\sqrt{x^2+9}}$ for $x=4{\cdot}2$.

Ans. 0·8144.

18. Find the approximate value of

$$\sqrt{\frac{x^2-x+1}{x^2+x+1}}$$

for $x=0{\cdot}3$.

Ans. 0·7.

19. Knowing that log 200=2·30103, find log 200·4. Compare the result with the tabulated figure. (Modulus $\mu=\log_e=0{\cdot}43429$).

Ans. log 200·4 = 2·30190. This is the same as the value in five-figure tables.

Hint. We write $y=\log x$. Then

$$\log(x+\Delta x)=y+\Delta y \approx y+dy=\log x+d(\log x) = \log x+\frac{1}{x}\log e\,dx$$

Show that for small absolute values of h:

20. $e^h \approx 1+h$. 21. $\tan h \approx h$.

22. $\arcsin h \approx h$. 23. $\arctan 2h \approx 2h$.

24. $\ln(1+\sqrt{h}) \approx \sqrt{h}$.

25. Show that for small absolute values of α,

$$\ln(1+\sin\alpha) \approx \alpha.$$

On §§ 74, 75.

26. Find the curvature and radius of curvature of the parabola $y^2=4x$ at the point (4, 4).

Ans. $-\dfrac{1}{10\sqrt{5}}$, $-10\sqrt{5}$.

27. Find the curvature and radius of curvature of the rectangular hyperbola $xy=12$ at the point (4, 3).

Ans. $\dfrac{24}{125}$ and $\dfrac{125}{24}$.

28. Find the radius of curvature of the curve $y=x^4-4x^3-18x^2$ at the point (0, 0).

Ans. $R=-\dfrac{1}{36}$.

29. Find the radius of curvature of the hyperbola $b^2x^2-a^2y^2=a^2b^2$, (*i*) at an arbitrary point, (*ii*) at the point $(a, 0)$.

Ans. (*i*) $R=-\dfrac{(b^4x^2+a^4y^2)^{\frac{3}{2}}}{a^4b^4}$;

(*ii*) $R=-\dfrac{b^2}{a}$.

30. At what point has the curve $y=e^x$ the least radius of curvature?

Ans. $\left(-\frac{1}{2}\ln 2, \frac{\sqrt{2}}{2}\right)$.

31. At what point is the absolute value of the radius of curvature of the curve $y=\ln x$ a minimum?

Ans. $\left(\frac{\sqrt{2}}{2}, -\frac{1}{2}\ln 2\right)$.

32. Show that the radius of curvature at any point (x, y) of the curve $x^{2/3}+y^{2/3}=a^{2/3}$ is three times the length of the perpendicular from the origin to the tangent to the curve at (x, y) (i.e. at the point where the radius of curvature is taken).

CHAPTER 9

INDEFINITE INTEGRALS

§ 77. Finding a function given its derivative or differential. Examples from mechanics and geometry. 1. We discovered in the differential calculus how to find the instantaneous velocity of a moving particle. We started from the law of motion $s = F(t)$ defining the variation of the path s in the course of time t and obtained the velocity v as the derivative of s with respect to t, i.e.

$$v = s' = F'(t)$$

(see § 40). Often, however, we want to solve just the reverse problem, i.e. given the velocity $v = f(t)$, find the path traversed s. This means that we want to re-establish some law of motion $s = F(t)$ in accordance with a given velocity $v = f(t)$, or in other words, given the velocity $v = f(t)$, to re-establish the function $s = F(t)$ of which v is the derivative.

For instance, the variation in time of the velocity v of a material particle falling in space is given by the law $v = gt$. To find the distance travelled, we have to find the function s of which v is the derivative. If the distance is measured from the initial instant, i.e. $s = 0$ when $t = 0$, we already know from (§ 40) that

$$s = \frac{1}{2} gt^2.$$

And in fact, $s' = \left(\frac{1}{2}gt^2\right)' = gt$ and $s = 0$ at $t = 0$.

2. We now consider this problem. The slope of the tangent to some unknown curve at any point $M(x, y)$ of the curve is given by

$$k = 2x,$$

and the curve passes through the origin. We are required to find the equation of the curve.

We know that, if the curve is given by the equation $y=F(x)$, the slope of the tangent at the point with abscissa x is equal to the derivative $y'=F'(x)$. Thus we have by hypothesis, $F'(x)=2x$. This means that we want here to find the unknown function $f(x)$ from its given derivative $2x$. It may easily be seen that the required function is $y=x^2$, since $y'=2x$, and in addition, the parabola $y=x^2$ passes through the origin.

3. From the purely mathematical point of view, the problem in both the above examples is as follows: We were given the derivative $f(x)$ of a certain function (the velocity in the first case, the slope of the tangent in the second). We had to find the function $y=F(x)$ such that its derivative is $f(x)$ (we wanted to discover the law of motion in the first case, and the equation of the curve in the second).

In other words, we had to solve in both cases a problem *the inverse of differentiation.*

In practice, the function $F(x)$ is given in terms of its given differential $f(x)\,dx$ rather than of its derivative $f(x)$. These are equivalent since

$$dF(x)=F'(x)\,dx,$$

so that

$$F'(x)\,dx=f(x)\,dx,$$

whence

$$F'(x)=f(x).$$

We have returned in essence to the problem of finding the function $F(x)$ of which $f(x)$ is the derivative.

A great many scientific and engineering problems are solved with the aid of the inverse operation to differentiation. We shall therefore make a detailed study of this operation, which is known as *integration.*

§ 78. Indefinite integrals. DEFINITION. *The function $F(x)$ is said to be a primitive of the function $f(x)$ if $f(x)$ is the derivative of $F(x)$, or what amounts to the same thing, if $f(x)\,dx$ is the differential of $F(x)$.* We clearly have here the relationship

$$F'(x)=f(x),$$

or $$dF(x)=f(x)\,dx.$$

For instance, sin x a primitive of cos x, since

$$(\sin x)'=\cos x,$$

or $$d(\sin x)=\cos x\,dx.$$

A primitive of x^2 is $\frac{x^3}{3}$, since

$$\left(\frac{x^3}{3}\right)'=x^2,$$

or $$d\left(\frac{x^3}{3}\right)=x^2\,dx.$$

We now notice that, as well as sin x, the functions $\sin x+1$, $\sin x-\sqrt{2}$, $\sin x+\pi$, and, in general,

$$\sin x+C,$$

where C is an arbitrary constant, are also primitives of cos x. For all these functions differ only in their constant terms. And since the derivative of a constant is zero, we have

$$(\sin x+1)'=\cos x\,,\ (\sin x-\sqrt{2t})'=\cos x,\ (\sin x+\pi)'=\cos x\,,$$

and, in general, $$(\sin x+C)'=\cos x.$$

Similarly, every function of the form $\frac{x^3}{3}+C$ is a primitive of x^2, because

$$\left(\frac{x^3}{3}+C\right)'=x^2.$$

It is shown in the more comprehensive courses of analysis that every function $f(x)$, continuous in a certain domain, has a primitive $F(x)$. It easily follows from this that $f(x)$ must in fact have an infinite set of primitives, differing from each other by a constant term. For, if

$$F'(x)=f(x),$$

then we also have

$$[F(x)+1]'=f(x) \quad \text{and} \quad [F(x)+3]'=f(x)$$

and
$$\left[F(x)-\frac{\sqrt{2}}{2}\right]'=f(x)$$

and, in general, $$[F(x)+C]'=f(x),$$

where C is an arbitrary constant.

Now the question naturally arises: if $F(x)$ is any given primitive of $f(x)$, continuous in the interval (a, b), does the expression $F(x)+C$ embrace all the possible primitives of $f(x)$? Is it possible for $f(x)$ to have other primitives that are not obtained from the expression $F(x)+C$, no matter what the value of the constant C? The answer is, in fact, that there are no such further primitives, and that $F(x)+C$ expresses the most general form of function having the derivative $f(x)$. The reasons for this situation may be seen from the following:

Let $F(x)$ and $F_1(x)$ be any two primitives of $f(x)$, which is defined in a certain domain. We form the difference

$$F_1(x)-F(x)=\varphi(x).$$

Since $F'(x)=f(x)$ and $F'_1(x)=f(x)$ for any x, we have

$$\varphi'(x)=F'_1(x)-F'(x)=f(x)-f(x)=0.$$

But the function $y=\varphi(x)$, whose derivative $y'=\varphi'(x)$ is zero in the interval (a, b), must be represented by a curve whose tangent is parallel to Ox at every point. Clearly, the only curve satisfying this requirement is a straight line parallel to Ox, so that the function $\varphi(x)$ must preserve a constant value throughout (a, b). For, if we suppose that the ordinates representing the values of $\varphi(x)$ for different values of x have different lengths, the curve $y=\varphi(x)$ must be rising or falling over portions of the interval, and its tangent cannot always be parallel to Ox *.

---

* This is proved rigorously in comprehensive courses of analysis by Lagrange's theorem. Our present arguments serve to explain the situation but essentially depend for support on geometrical intuition.

Therefore the difference $F_1(x) - F(x) = \phi(x)$ is constant, i.e.

$$F_1(x) - F(x) = C_1,$$

where C_1 is some number.

If, then, any two primitives $F_1(x)$ and $F(x)$ of the same function $f(x)$ differ only by a constant, having found one of them, say $F(x)$, we can obtain any other by adding some number C to $F(x)$. Hence the expression

$$F(x) + C,$$

in which C is regarded as an arbitrary constant, represents the set of all primitives of the function $f(x)$. In other words, $F(x) + C$ is the most general expression for a primitive of $f(x)$.

DEFINITION. *If $F(x)$ is any primitive of the function $f(x)$, i.e. $F'(x) = f(x)$, the expression $F(x) + C$, where C is an arbitrary constant, is termed the indefinite integral of $f(x)$* and is written symbolically as

$$\int f(x)\, dx.$$

The product $f(x)\,dx$ is called the *expression under the integral* whilst $f(x)$ is the *integrand.*

Taking Cartesian co-ordinates on the xOy plane, the equation

$$y = F(x) + C$$

defines a certain curve for a fixed value of C. With different values of C we obviously get different curves corresponding to different primitives. Hence the indefinite integral of $f(x)$ is said to represent a family of curves on the xOy plane depending on the parameter C. These curves are called *integral curves* of the function $f(x)$.

Example 1. Let $f(x) = x^3$. Then the indefinite integral is easily seen to be

$$\int x^3\, dx = \frac{x^4}{4} + C.$$

This is readily proved by carrying out the inverse operation, i.e. differentiation, for

$$\left(\frac{x^4}{4} + C\right)' = \frac{4x^3}{4} = x^3.$$

Example 2. If $f(x)=\frac{1}{\sqrt{x}}$, we have

$$\int \frac{1}{\sqrt{x}}\,dx=\int \frac{dx}{\sqrt{x}}=2\sqrt{x}+C.$$

This is also readily proved by differentiation, for

$$(2\sqrt{x}+C)'=2(\sqrt{x})'=\frac{2}{2\sqrt{x}}=\frac{1}{\sqrt{x}}.$$

We turn our attention to the fact that the differential and not the derivative of the required primitive is written under the integral sign $\left(\text{in our examples: } x^3dx \text{ and not } x^3;\ \frac{1}{\sqrt{x}}dx \text{ and not } \frac{1}{\sqrt{x}}\right)$. The reasons for this are partly historical and partly the greater convenience when dealing with complicated integrals.

The operation of finding the indefinite integral of a function is known as *integration.*

The following properties follow at once from the definition of indefinite integral:

1. $$(\int f(x)\,dx)'=f(x),$$

i.e. the derivative of the indefinite integral is the integrand.

2. $$d\,\{\int f(x)\,dx\}=f(x)\,dx,$$

i.e. *the differential of the indefinite integral is equal to the expression under the integral sign.*

3. Since $F(x)$ is a primitive of $F'(x)$, we have

$$\int F'(x)\,dx=F(x)+C,$$

or

$$\int dF(x)=F(x)+C,$$

i.e. *the integral of the differential of a function $F(x)$ is equal to $F(x)$ plus an arbitrary constant C.*

§ 79. Determining the arbitrary constant of integration from initial values of the variables. 1. We return to the problem of mechanics considered at the beginning of § 77. We considered there finding the law of motion of a particle falling in space, given its velocity $v=gt$ and the fact that the distance s traversed is measured from the initial instant (i.e. $s=0$ with $t=0$).

Since $v'=s$, we can now write the solution with the aid of the indefinite integral as

$$s=\int gt\,dt$$

whence, as is easily seen, $s=\frac{1}{2}gt^2+C.$ (1)

We have obtained an expression for s in which, apart from time t, there appears the arbitrary constant C. With the same fixed instant t, different values of C give us different values for the distance traversed. We were given the condition, however, that $s=0$ for $t=0$. Hence the constant C cannot be arbitrary but must have the value for which this condition is fulfilled. To find the required value of C, we substitute $s=0$ and $t=0$ in equation (1). We obtain

$$0=\frac{1}{2}g\cdot 0+C,$$

whence $C=0$, so that the equation for s is now fully defined as

$$s=\frac{1}{2}gt^2.$$

The values $t=0$ and $s=0$ quoted above are known as the *initial values* of t and s. Different initial values of t and s lead to different values of the constant C. If we were told, for instance, that the particle has already travelled a distance of 3 m at the instant $t=0$ of starting to measure time, we should have from equation (1)

$$3=\frac{1}{2}g\cdot 0+C.$$

whence $C=3$ and the equation of motion becomes

$$s=\frac{1}{2}gt^2+3.$$

2. We shall also reconsider the geometrical problem of § 77. Here we wanted to find the equation $y=F(x)$ of a curve such that the slope of its tangent k varies in accordance with the law

$$k=y'=2x$$

and such that the curve passes through the origin.
Since

$$y'=2x,$$

we have

$$y=\int 2x\,dx$$

or

$$y=x^2+C. \tag{2}$$

Each value of the constant C gives a corresponding determinate equation of the form (2) and therefore a definite curve (parabola). All these curves can be obtained from any one of them simply by displacing it parallel to Oy (fig. 88). The tangents to all the curves at points with the same abscissa x are parallel to each other, since for any C

$$\tan\alpha=k=(x^2+C)'=2x.$$

A unique parabola is defined from all these by the condition that the curve should pass through the origin (i.e. the initial values $x=0$, $y=0$ are the abscissa and ordinate of a point of the curve). For on substituting the initial values $x=0$, $y=0$ in equation (2), we obtain the value $C=0$ and at the same time a determinate equation for the required curve, i.e.

$$y=x^2.$$

If we were asked to find the parabola passing say through the point $(2, 8)$ instead of through the origin, we should have to substitute $x=2$ and $y=8$ in equation (2), thus obtaining $C=4$, so that the equation of the required parabola would be

$$y=x^2+4$$

3. We shall finally consider the problem of finding the function $y=F(x)$, whose derivative is x^2 and which has the value 12 at $x=1$.

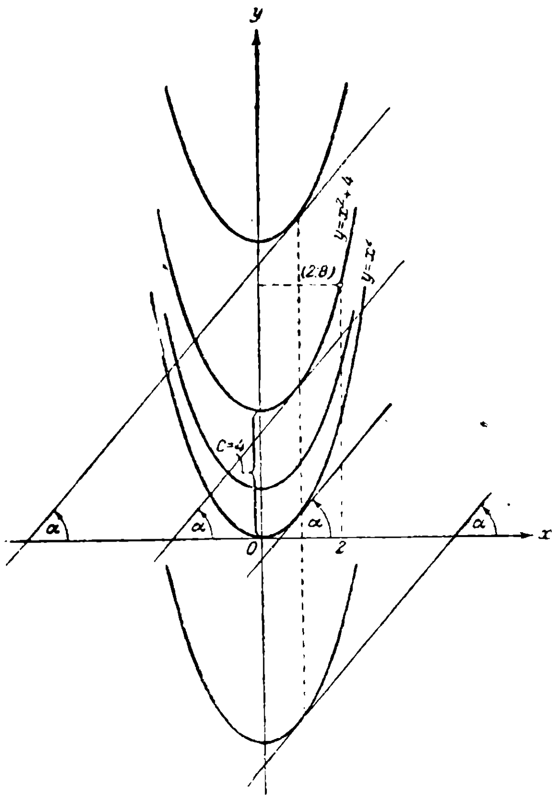


FIG. 88

It follows from these data that the required function is one of the primitives of x^2. On taking the indefinite integral of x^2, we obtain a whole family of primitives of x^2, i.e.

$$y=\int x^2\,dx=\frac{x^3}{3}+C.$$

To find the value of the constant C we use the initial values of the variables ($x=1$, $y=12$) to give

whence
$$12=\frac{1}{3}+C,$$

$$C=11\,\frac{2}{3}.$$

The required function is therefore

$$y=\frac{x^3}{3}+11\,\frac{2}{3}.$$

4. In all the above problems we have been required to find a definite primitive for a given function, given the initial values of the variables. If the initial values of the variables are not stated in the problem, we can only find the indefinite integral of the function, i.e. an entire family of its primitives $F(x)+C$, dependent on the single parameter C. Thus in the absence of initial values for the variables we cannot completely define the primitive of the given function. This is why the integral used in solving such problems is called an indefinite integral.

§ 80. Inversion of differentiation formulae (basic formulae for integration). Two rules for integration. When considering the basic properties of the indefinite integral that follow directly from its definition, we noted the extremely important equation

$$\int F'(x)\,dx=F(x)+C.$$

This equation enables us to find the integrals of some elementary functions by an inversion of familiar differentiation formulae. For we have by (3), e.g.

and since we know that $\int (\sin x)'\,dx=\sin x+C;$

it follows that $(\sin x)'=\cos x,$

Similarly, $\int \cos x\,dx=\sin x+C.$

$$\int\left(\frac{x^{n+1}}{n+1}\right)'dx=\frac{x^{n+1}}{n+1}+C \qquad (n\neq -1),$$

and since

$$\left(\frac{x^{n+1}}{n+1}\right)' = x^n,$$

we have

$$\int x^n\,dx = \frac{x^{n+1}}{n+1} + C \text{ and so on.}$$

All the integration formulae given below can be obtained in this way:

I. $\displaystyle\int x^a\,dx = \frac{x^{a+1}}{a+1} + C$ ($a \neq -1$; with $a = -1$ the formula becomes meaningless, for $a+1 = -1+1 = 0$, and division by zero is impossible).

II. $$\int \frac{dx}{x} = \int x^{-1}\,dx = \ln x + C.$$

III. $$\int a^x\,dx = \frac{a^x}{\ln a} + C.$$

IV. $$\int e^x\,dx = e^x + C.$$

V. $$\int \sin x\,dx = -\cos x + C.$$

VI. $$\int \cos x\,dx = \sin x + C.$$

VII. $$\int \frac{dx}{\cos^2 x} = \tan x + C.$$

VIII. $$\int \frac{dx}{\sin^2 x} = -\cot x + C.$$

IX. $$\int \frac{dx}{\sqrt{1-x^2}} = \operatorname{arc\,sin} x + C.$$

X. $$\int \frac{dx}{1+x^2} = \operatorname{arc\,tan} x + C.$$

The integration of other functions requires a knowledge of the basic rules and methods of the integral calculus, which we shall now discuss.

We mention first of all two fundamental rules for integration.

1. *A constant factor in the integrand can be taken outside the sign of integration*, i.e. if c is a constant (i.e. is independent of x), then

$$\int cf(x)\,dx = c\int f(x)\,dx.$$

Proof. By definition of integral we have

$$\left[\int cf(x)\,dx\right]' = cf(x), \quad \left[c\int f(x)\,dx\right]' = c\left[\int f(x)\,dx\right]' = cf(x).$$

Hence $$\int cf(x)\,dx = c\int f(x)\,dx.$$

Example.

$$\int \frac{\cos x}{5}\,dx = \int \frac{1}{5}\cos x\,dx = \frac{1}{5}\int \cos x\,dx = \frac{1}{5}\sin x + C.$$

2. *The integral of an algebraic sum of functions is equal to the algebraic sum of the integrals of the separate functions (and conversely).*

Confining ourselves to the sum of three functions, this property can be expressed as

$$\int [f(x)+\varphi(x)-\psi(x)]\,dx$$
$$= \int f(x)\,dx + \int \varphi(x)\,dx - \int \psi(x)\,dx.$$

Proof. On taking the derivatives of both sides of this equation we obtain

$$\left\{\int [f(x)+\varphi(x)-\psi(x)]\,dx\right\}' = f(x)+\varphi(x)-\psi(x);$$

$$\left[\int f(x)\,dx + \int \varphi(x)\,dx - \int \psi(x)\,dx\right]'$$
$$= \left[\int f(x)\,dx\right]' + \left[\int \varphi(x)\,dx\right]' - \left[\int \psi(x)\,dx\right]'$$
$$= f(x)+\varphi(x)-\psi(x).$$

Hence $$\int [f(x)+\varphi(x)-\psi(x)]\,dx$$
$$= \int f(x)\,dx + \int \varphi(x)\,dx - \int \psi(x)\,dx.$$

Example.

$$\int\left(\cos x - a^x + \frac{1}{\cos^2 x}\right)dx = \int \cos x\,dx - \int a^x\,dx +$$

$$+\int\frac{dx}{\cos^2 x}=\sin x-\frac{a^x}{\ln a}+\tan x+C$$

(by rule 2 and formulae VI, III and VII).

We do not write down the arbitrary constant each time when integrating the individual terms since the sum of several arbitrary constants can obviously be represented by the single letter C.

§ 81. Elementary methods of integration. 1. *Direct integration.* Methods of direct integration mean those in which the given integral can be reduced to one or more tabulated integrals. by using the two elementary rules for integration given in the previous article and/or by applying elementary transformations to the integrand. The above remarks will be made clear by examples.

1) $$\int(x^3+5x^2-7x+3)\,dx=?$$

Solution.

$$\int(x^3+5x^2-7x+3)\,dx$$

$$=\int x^3\,dx+5\int x^2\,dx-7\int x\,dx+3\int dx \quad \text{(by rules 2 and 1)}$$

$$=\frac{x^4}{4}+\frac{5x^3}{3}-\frac{7x^2}{2}+3x+C \qquad \text{(by formula I).}$$

2) $$\int\left(ax^2+bx+\frac{c}{x}+\frac{e}{x^2}\right)dx$$

$$=a\int x^2\,dx+b\int x\,dx+c\int\frac{dx}{x}+e\int x^{-2}\,dx \quad \text{(by rules 2 and 1)}$$

$$=a\frac{x^3}{3}+b\frac{x^2}{2}+c\ln x+e\frac{x^{-1}}{-1}+C \quad \text{(by formulae I and II)}$$

$$=\frac{1}{3}ax^3+\frac{1}{2}bx^2+c\ln x-\frac{e}{x}+C.$$

3) $$\int\left(\frac{2a}{\sqrt{x}}-\frac{b}{x^2}+3c\sqrt[3]{x^2}\right)dx$$

$$=2a\int x^{-\frac{1}{2}}\,dx-b\int x^{-2}\,dx+3c\int x^{\frac{2}{3}}\,dx$$

$$=2a\frac{x^{\frac{1}{2}}}{\frac{1}{2}}-b\frac{x^{-1}}{-1}+3c\frac{x^{\frac{5}{3}}}{\frac{5}{3}}+C \qquad \text{(by formula I)}$$

$$=4a\sqrt{x}+\frac{b}{x}+\frac{9}{5}cx^{\frac{5}{3}}+C.$$

4) $\int(x^2+2)\,x\,dx=\int(x^3+2x)\,dx$

$$=\int x^3\,dx+2\int x\,dx=\frac{x^4}{4}+x^2+C.$$

5) $\displaystyle\int\frac{x^3-3x^2+1}{x}\,dx=\int\left(x^2-3x+\frac{1}{x}\right)dx$

$$=\int x^2\,dx-3\int x\,dx+\int\frac{dx}{x}=\frac{x^3}{3}-\frac{3}{2}x^2+\ln x+C.$$

6) $\displaystyle\int\frac{1+2x^2}{x^2(1+x^2)}\,dx=\int\frac{(1+x^2)+x^2}{x^2(1+x^2)}\,dx$

$$=\int\frac{1+x^2}{x^2(1+x^2)}\,dx+\int\frac{x^2}{x^2(1+x^2)}\,dx$$

$$=\int\frac{dx}{x^2}+\int\frac{dx}{1+x^2}$$

$$=\int x^{-2}\,dx+\arctan x=-\frac{1}{x}+\arctan x+C.$$

2. *Integration by substitution.* The substitution method is perhaps the most important method of integration. It will be explained with the aid of several examples.

1) Let the required integral be

$$\int\sqrt{x+1}\,dx.$$

If the integral given were of the form

$$\int\sqrt{x}\,dx,$$

we could evaluate it by using formula I $\left(a=\frac{1}{2}\right)$. Hence it seems reasonable to bring in the auxiliary variable $t=x+1$ (or in more conventional language, make the substitution $x+1=t$) in order to reduce the integrand to the form $\sqrt{t}$. But to write the full expression under the integral in terms of t, we need to express dx as well as $\sqrt{x+1}$ in terms of t and dt. We do this by taking the differentials of both sides of the equation $x+1=t$, whence we obtain

$$d(x+1)=dt$$

or

$$dx=dt.$$

We now have

$$\int\sqrt{x+1}\,dx=\int\sqrt{t}\,dt=\int t^{\frac{1}{2}}\,dt=\frac{2}{3}\,t^{\frac{3}{2}}+C.$$

To express the result obtained in terms of the old variable x we only need to replace t by $x+1$. Thus we finally obtain

$$\int\sqrt{x+1}\,dx=\frac{2}{3}\,(x+1)^{\frac{3}{2}}+C.$$

2) Let us evaluate $\int\sqrt{3x+1}\,dx$. We use the same arguments as in example 1 and make the substitution $3x+1=t$. Then $d(3x+{}$ $+1)=3dx=dt$, $dx=\frac{1}{3}dt$ and therefore

$$\int\sqrt{3x+1}\,dx=\int\sqrt{t}\left(\frac{1}{3}\,dt\right)=\frac{1}{3}\int\sqrt{t}\,dt$$

$$=\frac{1}{3}\cdot\frac{2}{3}\,t^{\frac{3}{2}}+C=\frac{2}{9}\,(3x+1)^{\frac{3}{2}}+C.$$

3) We consider $\int\cos 5x\,dx$. This recalls the tabulated integral $\int\cos x\,dx$. The natural substitution is therefore $5x=t$; then $5dx=dt$, $dx=\frac{dt}{5}$ and hence

$$\int \cos 5x\,dx=\frac{1}{5}\int \cos t\,dt=\frac{1}{5}\sin t+C=\frac{1}{5}\sin 5x+C.$$

4) We now turn to a more difficult example. Let the required integral be

$$\int \sin^2 x \cos x\,dx.$$

This resembles none of the tabulated integrals, so that the substitution needed to reduce it to one of the tabulated forms is more difficult to find than in the previous examples. The required substitution is based on the observation that the differential of $\sin x$ is $\cos x\,dx$. Thus if we put $\sin x=t$, we obtain $d(\sin x)=\cos x\,dx=dt$ and

$$\int \sin^2 x \cos x\,dx=\int t^2\,dt=\frac{t^3}{3}+C=\frac{\sin^3 x}{3}+C.$$

5) The same sort of arguments enable us to find $\int \ln x\frac{dx}{x}$.

We note that $(\ln x)'=\frac{1}{x}$ and make the substitution $\ln x=t$; then $d(\ln x)=\frac{dx}{x}=dt$ and

$$\int \ln x\frac{dx}{x}=\int t\,dt=\frac{t^2}{2}+C=\frac{1}{2}\ln^2 x+C.$$

6) Let the required integral be $\int \frac{\sin x\,dx}{\sqrt[3]{\cos^2 x}}$. We put $\cos x=t$; then $d(\cos x)=-\sin x\,dx=dt$, $\sin x\,dx=-dt$. Therefore

$$\int \frac{\sin x\,dx}{\sqrt[3]{\cos^2 x}}=\int \cos^{-\frac{2}{3}} x\cdot\sin x\,dx=-\int t^{-\frac{2}{3}}\,dt$$

$$=-3t^{\frac{1}{3}}+C=-3\sqrt[3]{\cos x}+C.$$

It must be borne in mind when choosing the substitition that the differential of the function which is replaced by the new variable t must appear as a factor in the expression under the inte-

gral *; this factor gives the differential dt of the new variable, as a result of which the expression under the integral is simplified.

All the integrals in the examples of no. 2 of § 81 reduce to the same tabulated integral (I). In other words, the same tabulated formula allows a set of different integrals to be evaluated. We show how the same procedure can be applied to formula (II), i.e. to

$$\int \frac{dx}{x} = \ln x + C.$$

7) $\int \frac{dx}{x+1} = ?$ On substituting $x+1=t$, we obtain $dx=dt$, and consequently

$$\int \frac{dx}{x+1} = \int \frac{dt}{t} = \ln t + C = \ln (x+1) + C.$$

8) $\int \frac{x^2 dx}{x^3+2} = ?$ We put $x^3+2=t$; then $3x^2dx=dt, x^2dx=\frac{dt}{3}$.

Hence

$$\int \frac{x^2\,dx}{x^3+2} = \frac{1}{3}\int \frac{dt}{t} = \frac{1}{3}\ln t + C = \frac{1}{3}\ln (x^3+2) + C.$$

9) $\int \tan x\,dx = \int \frac{\sin x\,dx}{\cos x}$. On substituting $\cos x = t$, we obtain $-\sin x\,dx = dt$, $\sin x\,dx = -dt$. Hence

$$\int \tan x\,dx = -\int \frac{dt}{t} = -\ln t + C = -\ln \cos x + C.$$

10) $\int \frac{dx}{x(1+\ln x)} = ?$ Substituting $1+\ln x = t$, we obtain

$$\frac{dx}{x} = dt,$$

* If not strictly the differential, then the differential multiplied by a constant, cf. examples 2, 3, 6.

whence $\int \frac{dx}{x(1+\ln x)} = \int \frac{dt}{t} = \ln t + C = \ln(1+\ln x) + C.$

The integrals below reduce to the formula

$$\int \frac{dt}{1+t^2} = \text{arc tan } t + C.$$

11) $\int \frac{\cos x\, dx}{1+\sin^2 x} = ?$ We put $\sin x = t$; then $\cos x\, dx = dt$.

Hence $\int \frac{\cos x\, dx}{1+\sin^2 x} = \int \frac{dt}{1+t^2} = \text{arc tan } t + c = \text{arc tan} \sin x + C.$

12) $\int \frac{2x^3\, dx}{1+x^8} = ?$ Substituting $x^4 = t$, we obtain

$4x^3\, dx = dt$, $2x^3\, dx = \frac{1}{2}\, dt$ and

$$\int \frac{2x^3\, dx}{1+x^8} = \frac{1}{2} \int \frac{dt}{1+t^2} = \frac{1}{2} \text{ arc tan } t + C = \frac{1}{2} \text{arc tan } x^4 + C.$$

13) $\int \frac{5x^3\, dx}{4+x^8} = ?$ We first reduce this integral to the form

$$\int \frac{5x^3\, dx}{4+x^8} = \frac{5}{4} \int \frac{x^3\, dx}{1+\left(\frac{x^4}{2}\right)^2};$$

and then put $\frac{1}{2} x^4 = t$. We have $2x^3 dx = dt$, $x^3 dx = \frac{1}{2}\, dt$, and so

$$\int \frac{5x^3\, dx}{4+x^8} = \frac{5}{4} \int \frac{dt}{2(1+t^2)} = \frac{5}{8} \int \frac{dt}{1+t^2}$$

$$= \frac{5}{8} \text{ arc tan } t + C = \frac{5}{8} \text{ arc tan } \frac{1}{2} x^4 + C.$$

The following are some examples of integrals reducing to various tabulated integrals with the aid of suitable substitutions.

14) $\int e^{2x}dx = ?$ We put $2x=t$; then $dx=\frac{1}{2}\,dt$. Hence

$$\int e^{2x}\,dx=\frac{1}{2}\int e^t\,dt=\frac{1}{2}\,e^t+C=\frac{1}{2}\,e^{2x}+C.$$

15) $\int e^{x^3+1}x^2dx=?$ Substituting $x^3+1=t$, we obtain $3x^2dx =dt$, $x^2dx=\frac{dt}{3}$. Therefore

$$\int e^{x^3+1}\,x^2\,dx=\frac{1}{3}\int e^t\,dt=\frac{1}{3}\,e^t+C=\frac{1}{3}\,e^{x^3+1}+C.$$

16) $\int(a^{nx}-e^{mx})\,dx=\int a^{nx}dx-\int e^{mx}dx$. We evaluate the first integral on the right-hand side by putting $nx=t$; then $n\,dx=dt$ and $dx=\frac{dt}{n}$.

The second integral is evaluated by substituting $mx=z$; then $dx=\frac{dz}{m}$. Hence

$$\int(a^{nx}-e^{mx})\,dx=\int a^t\,\frac{dt}{n}-\int e^z\,\frac{dz}{m}=\frac{1}{n}\,\frac{a^t}{\ln a}-\frac{1}{m}\,e^z+C$$

$$=\frac{a^{nx}}{n\ln a}-\frac{e^{mx}}{m}+C.$$

17) $\int\frac{\sin x\,dx}{a+b\cos x}=?$ We put $a+b\cos x=t$; then $-b\sin x\,dx =dt$ and $\sin x\,dx=-\frac{dt}{b}$. Hence

$$\int\frac{\sin x\,dx}{a+b\cos x}=-\frac{1}{b}\int\frac{dt}{t}=-\frac{1}{b}\ln(a+b\cos x)+C.$$

18) $\int\frac{x^2dx}{\cos^2x^3}=?$ We put $x^3=t$; then $x^2dx=\frac{dt}{3}$.

Hence

$$\int \frac{x^2\,dx}{\cos^2 x^3} = \frac{1}{3}\int \frac{dt}{\cos^2 t} = \frac{1}{3}\tan t + C = \frac{1}{3}\tan t + C \quad \text{(by formula VII)}$$

19) $$\sqrt{16-qx^2} = \frac{1}{3}\tan x^3 + C.$$

This integral reduces to integral IX:

$$\int \frac{dx}{\sqrt{16-9x^2}} = \frac{1}{4}\int \frac{dx}{\sqrt{1-\left(\frac{3}{4}x\right)^2}}.$$

We put $\frac{3}{4}x = t$; then $dx = \frac{4}{3}\,dt$. Hence

$$\int \frac{dx}{\sqrt{16-9x^2}} = \frac{1}{3}\int \frac{dt}{\sqrt{1-t^2}} = \frac{1}{3}\arcsin t = \frac{1}{3}\arcsin \frac{3}{4}x + C.$$

20) $$\int \cos^3 x\,dx = \int \cos^2 x \cos x\,dx = \int (1-\sin^2 x)\cos x\,dx$$
$$= \int \cos x\,dx - \int \sin^2 x \cos x\,dx.$$

The first of the last two integrals is tabulated. We find the second (cf. example 4) by putting $\sin x = t$; then $\cos x\,dx = dt$. Hence

$$\int \cos x\,dx - \int \sin^2 x \cos x\,dx = \sin x - \int t^2\,dt$$
$$= \sin x - \frac{t^3}{3} + C = \sin x - \frac{\sin^3 x}{3} + C.$$

In the examples of integration by the substitution method so far considered the new variable has always been a function of the old argument x. Sometimes, however, the reverse process is used for the substitution method, the old argument x being replaced by a suitably chosen function of the new variable t.

Suppose, for instance, that we want to evaluate $\int \frac{dx}{(1+x^2)^{3/2}}$.

We put $x=\tan t$; then $dx=\dfrac{dt}{\cos^2 t}$. Hence

$$\int\frac{dx}{(1+x^2)^{3/2}}=\int\frac{dt}{\cos^2 t\,(1+\tan^2 t)^{3/2}}=\int\frac{dt}{\cos^2 t\left(\dfrac{1}{\cos^2 t}\right)^{3/2}}$$

$$=\int\frac{\cos^3 t\,dt}{\cos^2 t}=\int\cos t\,dt=\sin t+C.$$

The result has to be expressed in terms of the old argument x. Since $x=\tan t$, we have

$$\sin t=\frac{\tan t}{\sqrt{1+\tan^2 t}}=\frac{x}{\sqrt{1+x^2}}.$$

Hence
$$\int\frac{dx}{(1+x^2)^{3/2}}=\frac{x}{\sqrt{1+x^2}}+C.$$

We take another example. Let the required integral be $\int\sqrt{a^2-x^2}\,dx$. We introduce a new variable t by putting $x=a\sin t$; then $dx=a\cos t\,dt$, and therefore

$$\int\sqrt{a^2-x^2}\,dx=\int\sqrt{a^2(1-\sin^2 t)}\cdot a\cos t\,dt$$
$$=a^2\int\cos^2 t\,dt.$$

The last integral is evaluated with the aid of the familiar trigonometric formula: $\cos^2 t=\dfrac{1}{2}(1+\cos 2t)$. Therefore

$$a^2\int\cos^2 t\,dt=\frac{a^2}{2}\int(1+\cos 2t)\,dt=\frac{a^2}{2}\left[\int dt+\int\cos 2t\,dt\right].$$

The first of the last two integrals is obtained directly:

$$\int dt=t+C_1$$

The second is obtained by substituting $2t=u$; then $2dt=du$, d $=\dfrac{1}{2}du$, and

$$\int \cos 2t\, dt = \frac{1}{2}\int \cos u\, du = \frac{1}{2}\sin u + C_2 = \frac{1}{2}\sin 2t + C_2.$$

Thus

$$\int \sqrt{a^2-x^2}\, dx = \frac{a^2}{2}\left[t + \frac{1}{2}\sin 2t\right] + C$$

$\left\{\text{we have simply written } C \text{ for } \frac{1}{2}a^2(C_1+C_2)\right\}$.

To return to the old argument x in the final result, we find from the relationship $x = a\sin t$:

$$\sin t = \frac{x}{a}, \qquad \cos t = \sqrt{1-\frac{x^2}{a^2}} = \frac{1}{a}\sqrt{a^2-x^2},$$

$$\sin 2t = 2\sin t\cos t = 2\cdot\frac{x}{a}\cdot\frac{1}{a}\sqrt{a^2-x^2},$$

$$t = \arcsin\frac{x}{a}.$$

Hence, we finally obtain

$$\int \sqrt{a^2-x^2}\, dx = \frac{a^2}{2}\arcsin\frac{x}{a} + \frac{x}{2}\sqrt{a^2-x^2} + C.$$

Exercises

On §§ 78, 81 *and* 79.

Evaluate the following integrals:

1. $\int x^6\, dx$. Ans. $\frac{x^7}{7} + C$.

2. $\int 4x^5\, dx$. Ans. $\frac{2}{3}x^6 + C$.

3. $\int (1-2x)\, dx$. Ans. $x - x^2 + C$.

4. $\int (ax+b)\, dx$. Ans. $\frac{a}{2}x^2 + bx + C$.

5. $\int (2-3u^4)\,du$. Ans. $2u-\frac{3}{5}u^5+C$.

6. $\int (3x^2-2x+1)\,dx$. Ans. x^3-x^2+x+C.

7. $\int (4x^3+3x^2+4x-3)\,dx$. Ans. $x^4+x^3+2x^2-3x+C$.

8. $\int (ax^3+bx^2+cx+e)\,dx$. Ans. $\frac{a}{4}x^4+\frac{b}{3}x^3+\frac{c}{2}x^2+$ $+ex+C$.

9. $\int \frac{dx}{x^2}$. Ans. $-\frac{1}{x}+C$.

10. $\int \frac{5dt}{t^3}$. Ans. $-\frac{5}{2t^2}+C$.

11. $\int\left(\frac{1}{x^4}-\frac{1}{x^3}\right)dx$. Ans. $-\frac{1}{3x^3}+\frac{1}{2x^2}+C$.

12. $\int\left(x^3-x^2+\frac{1}{x^2}-\frac{1}{x^3}\right)dx$. Ans. $\frac{x^4}{4}-\frac{x^3}{3}-\frac{1}{x}+\frac{1}{2x^2}+C$.

13. $\int \frac{x^5-x^3+1}{x^2}\,dx$. Ans. $\frac{x^4}{4}-\frac{x^2}{2}-\frac{1}{x}+C$.

14. $\int \sqrt[3]{t}\,dt$. Ans. $\frac{3}{4}t\sqrt[3]{t}+C$.

15. $\int 2\sqrt[5]{t^3}\,dt$. Ans. $\frac{5}{4}t\sqrt[5]{t^3}+C$.

16. $\int (2u-3\sqrt{u})\,du$. Ans. $u^2-2u\sqrt{u}+C$

17. $\int \frac{dv}{3\sqrt{v}}$. Ans. $\frac{2}{3}\sqrt{v}+C$.

18. $\int \frac{dx}{\sqrt[3]{x^2}}$. Ans. $3\sqrt[3]{x}+C$.

19. $\int \frac{dx}{\sqrt{2x}}$. Ans. $\sqrt{2x}+C$.

20. $\int \frac{x^2+\sqrt{x^3}+3}{\sqrt{x}}\,dx.$ Ans. $\frac{2}{5}x^{\frac{5}{2}}+\frac{x^2}{2}+6\sqrt{x}+C.$

21. $\int 7^x\,dx.$ Ans. $\frac{7^x}{\ln 7}+C.$

22. $\int (x^3+3^x)\,dx.$ Ans. $\frac{x^4}{4}+\frac{3^x}{\ln 3}+C.$

23. $\int (x^3-1)^2\,dx.$ Ans. $\frac{x^7}{7}-\frac{x^4}{2}+x+C.$

24. $\int \left(\frac{2+x}{x}\right)^2 dx.$ Ans. $-\frac{4}{x}+4\ln x+x+C.$

25. $\int \frac{(x+1)^2}{\sqrt{x}}\,dx.$ Ans. $\frac{2}{5}x^{\frac{5}{2}}+\frac{4}{3}x\sqrt{x}+$
$+2\sqrt{x}+C.$

26. $\int \frac{18x^2-2}{3x-1}\,dx.$ Ans. $3x^2+2x+C.$

27. $\int \frac{4-x}{2+\sqrt{x}}\,dx.$ Ans. $2x-\frac{2}{3}x\sqrt{x}+C.$

28. $\int \frac{(1+x)^2\,dx}{x(1+x^2)}.$ Ans. $\ln x+2\arctan x+C.$

29. $\int \frac{\sin 2x}{\sin x}\,dx.$ Ans. $2\sin x+C.$

30. $\int \frac{dx}{\sin^2 x\cos^2 x}.$ Ans. $\tan x-\cot x+C.$

31. $\int \cos 3x\,dx.$ Ans. $\frac{1}{3}\sin 3x+C.$

32. $\int \left(\sin\frac{x}{2}+\cos 3x\right)dx.$ Ans. $-2\cos\frac{x}{2}+\frac{1}{3}\sin 3x+C$

33. $\int (e^x + e^{-x})\, dx.$ Ans. $e^x - e^{-x} + C.$

34. $\int \frac{a}{2}\left(e^{\frac{x}{a}} + e^{-\frac{x}{a}}\right) dx.$ Ans. $\frac{a^2}{2}\left(e^{\frac{x}{a}} - e^{-\frac{x}{a}}\right) + C.$

35. $\int (x+1)^{\frac{3}{2}}\, dx.$ Ans. $\frac{2}{5}(x+1)^{\frac{5}{2}} + C.$

36. $\int (2+3x)^{\frac{3}{2}}\, dx.$ Ans. $\frac{2}{15}(2+3x)^{\frac{5}{2}} + C.$

37. $\int \frac{dx}{\sqrt{2x+1}}.$ Ans. $\sqrt{2x+1} + C.$

38. $\int \frac{x\, dx}{x^2+3}.$ Ans. $\frac{1}{2}\ln(x^2+3) + C.$

39. $\int \sqrt{2x^2+1}\, x\, dx.$ Ans. $\frac{1}{6}\sqrt{(2x^2+1)^3} + C.$

40. $\int \sqrt{a^2+b^2x^2}\, x\, dx.$ Ans. $\frac{1}{3b^2}(a^2+b^2x^2)^{\frac{3}{2}} + C.$

41. $\int \frac{x\, dx}{\sqrt{x^2-2}}.$ Ans. $\sqrt{x^2-2} + C$

42. $\int \sqrt{2+e^x}\, e^x\, dx.$ Ans. $\frac{2}{3}(2+e^x)^{\frac{3}{2}} + C.$

43. $\int \frac{x^2\, dx}{(a^2+x^3)^{\frac{1}{2}}}.$ Ans. $\frac{2}{3}\sqrt{a^2+x^3} + C.$

44. $\int \frac{e^{2x}\, dx}{e^{2x}+3}.$ Ans. $\ln\sqrt{e^{2x}+3} + C.$

45. $\int e^{x^2}\, x\, dx.$ Ans. $\frac{1}{2}e^{x^2} + C.$

46. $\int e^{-\frac{1}{x}} \frac{dx}{x^2}$. Ans. $e^{-\frac{1}{x}}+C$.

47. $\int e^{\cos x} \sin x\, dx$. Ans. $-e^{\cos x}+C$.

48. $\int \frac{dx}{x \ln x^2}$. Ans. $\frac{1}{2} \ln \ln x+C$.

49. $\int a^{x^3} x^2\, dx$. Ans. $\frac{1}{3} \frac{a^{x^3}}{\ln a}+C$.

50. $\int a^x e^x\, dx$. Ans. $\frac{a^x e^x}{1+\ln a}+C$.

51. $\int \cot x\, dx$. Ans. $\ln \sin x+C$.

52. $\int x^2 \sin 3x^3\, dx$. Ans. $-\frac{1}{9} \cos 3x^3+C$.

53. $\int x \cos (a+bx^2)\, dx$. Ans. $\frac{1}{2b} \sin (a+bx^2)+C$.

54. $\int \frac{5dx}{\cos^2 bx}$. Ans. $\frac{5}{b} \tan bx+C$.

55. $\int \frac{dx}{\cos^2 (a-bx)}$. Ans. $-\frac{1}{b} \tan (a-bx)+C$.

56. $\int \frac{dx}{\sin^2 \frac{x}{n}}$. Ans. $-n \cot \frac{x}{n}+C$.

57. $\int \frac{dx}{3+4x^2}$. Ans. $\frac{\sqrt{3}}{6} \text{arc tan} \frac{2x}{\sqrt{3}}+C$.

58 $\int \frac{dx}{\sqrt{25-9x^2}}$. Ans. $\frac{1}{3} \text{arc sin} \frac{3x}{5}+C$.

59. $\int \frac{7dx}{\sqrt{3-5x^2}}$. Ans. $\frac{7}{\sqrt{5}} \arcsin \sqrt{\frac{5}{3}}\, x + C$.

60. $\int \frac{\cos x\, dx}{a^2+\sin^2 x}$. Ans. $\frac{1}{a} \arctan\left(\frac{\sin x}{a}\right) + C$.

61. $\int \frac{e^x\, dx}{\sqrt{1-e^{2x}}}$. Ans. $\arcsin e^x + C$.

62. $\int \frac{dx}{x\sqrt{1-\ln^2 x}}$. Ans. $\arcsin(\ln x) + C$.

63. $\int \sin^3 x \cos x\, dx$. Ans. $\frac{\sin^4 x}{4} + C$.

64. $\int \frac{dx}{x\sqrt{x^2-1}}$ $\left(\text{substitute } x = \frac{1}{z}\right)$. Ans. $-\arcsin \frac{1}{x} + C$.

65. $\int \frac{x^2\, dx}{\sqrt{a^2-x^2}}$ (substitute $x = a \sin t$).

Ans. $\frac{a^2}{2} \arcsin \frac{x}{a} - \frac{x}{2}\sqrt{a^2-x^2} + C$.

66. $\int \frac{x^2\, dx}{(x^2+1)^2}$ (substitute $x = \tan t$).

Ans. $\frac{1}{2}\left(\arctan x - \frac{x}{x^2+1}\right) + C$.

67. Find the function whose derivative is $x-3$, given that the function has the value 9 at $x=2$.

Ans. $\frac{x^2}{2} - 3x + 13$.

68. Find the function whose derivative is $3+x-5x^2$, given that its value at $x=6$ is -200.

Ans. $124+3x+\frac{x^2}{2}-\frac{5x^3}{3}$.

69. Find the function whose derivative is $\sin x+\cos x$, given that its value at $x=\frac{1}{2}\pi$ is 2.

Ans. $\sin x-\cos x+1$.

70. Find the equation of the curve that passes through the point $(0, 1)$, given that the slope of its tangent is equal to x.

Ans. $y=\frac{x^2}{2}+1$.

71. Find the equation of the curve which passes through the point $(1, -5)$, the slope k of the tangent to the curve at any point being given by the expression $k=1-x$.

Ans. $y=x-\frac{x^2}{2}-\frac{11}{2}$.

72. Find the equation of the curve which passes through the point $(0, a)$, the slope k of the tangent to the curve at any point being given by the expression $k=\frac{1}{2}(e^{x/a}-e^{-x/a})$.

Ans. $y=\frac{a}{2}\left(e^{\frac{x}{a}}+e^{-\frac{x}{a}}\right)$.

73. The velocity v of a body is given by $v=(3t^2+4t)$ m/sec. After $t=2$ sec the body has traversed a distance $s=16$ m. Find the law of motion of the body.

Ans. $s=t^3+2t^2$.

Hint. The velocity v is the derivative of the distance s with respect to time.

74. The velocity v of a body is given by $v=2\cos t$ m/sec. After $t=\pi/4$ sec. the body as traversed a distance s equal to 10 m. Find the law of motion.

Ans. $s=2\sin t+10-\sqrt{2}$.

75. The acceleration j of a body is given by $j=(t^2+1)$ m/sec^2. At the instant $t=0$ the velocity $v=1$ m/sec and the distance $s=0$. Find the law of motion.

Ans. $s=\frac{t^4}{12}+\frac{t^2}{2}+t$.

CHAPTER 10

THE DEFINITE INTEGRAL

§ 82. Definite integral as an area. Evaluation of the definite integral with the aid of the indefinite integral.

1. Let the function $y=f(x)$ be continuous and non-negative in the interval (a, b), and let A and B denote the points of the graph of the function corresponding to $x=a$ and $x=b$ (fig. 89). We consider the area $aABb$, bounded at the top by the arc AB of the curve $y=f(x)$, at the bottom by the segment (a, b) of the Ox axis, and at the sides by the perpendiculars Aa and Bb from

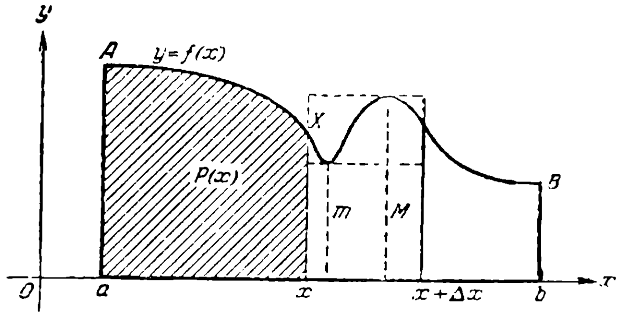


FIG. 89

the ends of arc AB to Ox*. This form of area is known as a curvilinear trapezium. We shall consider the problem of determining the area P of the curvilinear trapezium $aABb$.

We consider first the area $aAXx$, cut off from $aABb$ by the perpendicular xX drawn to the Ox axis at an arbitrary point x of (a, b). As x varies, both the shape and area of $aAXx$ vary, so

* We economize on symbols by writing a and b here for the points of Ox corresponding to $x=a$ and $x=b$. If $f(a)=0$ or $f(b)=0$ (or $f(a)=f(b)=0$), the distance aA or bB (or both) reduces to a point.

that the area of $aAXx$ must be a function of the argument x, which we shall denote by $P(x)$.

We now show that the function $P(x)$ has a derivative $P'(x)$, such that $P'(x)=f(x)$.

We give the increment $\Delta x>0$ to the abscissa x of the point X, in which case the area $P(x)$ receives the increment $\Delta P(x)$. We write m and M respectively for the least and greatest values of $f(x)$ in the interval $[x, x+\Delta x]$ * and compare the area $\Delta P(x)$ with the areas of the rectangles on the same base Δx but with heights m and M. We clearly have

$$m\,\Delta x < \Delta P(x) < M\,\Delta x,$$

whence

$$m < \frac{\Delta P(x)}{\Delta x} < M. \tag{1}$$

If we now let Δx tend to zero we shall have by virtue of the continuity of $f(x)$ (§ 38)

$$\lim_{\Delta x \to 0} m = \lim_{\Delta x \to 0} M = f(x).$$

(It is clear from the figure that, as $\Delta x \to 0$, the ordinates representing the least value m and greatest value M of $f(x)$ in the interval Δx will be moved from right to left and will tend to the common limit $xX=f(x)$.) It now follows from (1) that we also have

$$\lim_{\Delta x \to 0} \frac{\Delta P(x)}{\Delta x} = f(x).$$

But the limit of the ratio of the increment $\Delta P(x)$ of the function $P(x)$ to the increment Δx of its argument x is, by definition, the derivative of $P(x)$, i.e.

$$P'(x)=f(x).$$

If $\Delta x<0$, we also have $\Delta P(x)<0$, so that

$$m\,\Delta x > \Delta P(x) > M\Delta x,$$

---

* We omit the proof that a function continuous in a closed interval has a least value m and greatest value M in the interval. A proof of this property of continuous functions will be found in any comprehensive course on analysis.

whence

$$m < \frac{\Delta P(x)}{\Delta x} < M.$$

Since $\lim\limits_{\Delta x \to 0} m = \lim\limits_{\Delta x \to 0} M = f(x)$, similar arguments to the above lead us to the same result, viz.

$$\lim_{\Delta x \to 0} \frac{\Delta P(x)}{\Delta x} = f(x),$$

i.e.

$$P'(x) = f(x).$$

We have thus established the following theorem:

THEOREM. *The derivative of the variable area $P(x)$ of the curvilinear trapezium $aAXx$ is equal to the ordinate $y=f(x)$.*

2. It follows from the theorem just proved that the variable area $P(x)$ is one of the primitives of the function $f(x)$ This primitive is distinguished from all the others by the fact that it vanishes at $x=a$, for with $x=a$ the ordinate xX coincides with aA, the trapezium $aAXx$ degenerates into the straight line aA, and its area becomes zero.

Thus, given any primitive $F(x)$ (no matter which) of $f(x)$, we have (see § 78)

$$P(x) = F(x) + C_0,$$

where C_0 is a number which is readily determined by putting $x=a$ in the last equation. This gives us

$$0 = F(a) + C_0,$$

whence

$$C_0 = -F(a).$$

Hence, no matter what the primitive $F(x)$ of $f(x)$ and no matter what the value of x in the interval $[a, b]$, we have

$$P(x) = F(x) - F(a). \tag{2}$$

Example. Let us find the area $P(x)$ of the curvilinear trapezium bounded by the parabola $y=x^2$, the straight line $x=1$, the intercept $[1, x]$ of the Ox axis, and the variable ordinate xX of a point of the parabola (fig. 90).

Solution. One of the primitives of x^2 is $\frac{x^3}{3}$, i.e.

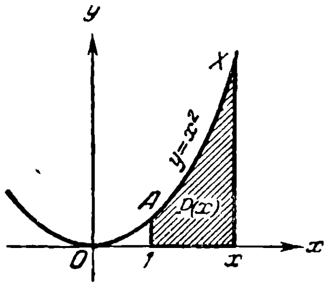


FIG. 90

$$P(x)=\frac{x^3}{3}+C_0.$$

Putting $x=1$ and observing that $P(1)=0$, we obtain

$$0=\frac{1}{3}+C_0,$$

whence

$$C_0=-\frac{1}{3};$$

so that

$$P(x)=\frac{x^3}{3}-\frac{1}{3}.$$

3. We return to fig. 89. We found that, if $F(x)$ is any primitive of $f(x)$, the area $P(x)$ of the trapezium $aAXx$ is given by equation (2), i.e.

$$P(x)=F(x)-F(a).$$

It now becomes clear that the area of the curvilinear trapezium $aABb$ based on the intercept $[a, b]$ is a number equal to the value of $P(x)$ at $x=b$, i.e. the number $P(b)$. We find from equation (2):

$$P(b)=F(b)-F(a). \tag{2*}$$

It follows from the above that we can find the area of the curvilinear trapezium bounded by the curve $y=f(x)$, the straight lines $x=a$, $x=b$ and the intercept $[a, b]$ of Ox simply by finding any *primitive $F(x)$ of $f(x)$ and subtracting its value at $x=a$ from its value at $x=b$,* i.e. the area is given by $F(b)-F(a)$.

We notice the fact that the area of the trapezium is the value $P(b)$ of a completely defined primitive $P(x)$ of $f(x)$, i.e. the primitive which vanishes with $x=a (P(a)=0)$. It is quite unnecessary to know the function $P(x)$, however, in order to find $P(b)$, which is simply obtained by subtracting the value of *any* primitive of $f(x)$ at $x=a$ from its value at $x=b$.

We take as an example the area bounded by the parabola $y=x^2$, the ordinates $x=1$, $x=3$ and the interval $[1, 3]$ of the Ox axis. Since one of the primitives of x^2 is $\frac{x^3}{3}$, on subtracting its value at $x=1$ from its value at $x=3$, we obtain

$$9-\frac{1}{3}=8\frac{2}{3}.$$

The required area is thus equal to $8\frac{2}{3}$ sq. units.

We now take any other primitive of x^2, say $\frac{x^3}{3}+5$. The difference between its values at $x=3$ and $x=1$ is

$$(9+5)-\left(\frac{1}{3}+5\right)=8\frac{2}{3},$$

i.e. we arrive at the same result as above.

4. We now extend the term "curvilinear trapezium" to include any figure bounded by the arc of a curve $y=f(x)$, the ordinates $x=a$, $x=b$ of the curve, and the intercept $[a, b]$ of Ox, no matter what the values of $f(x)$ in $[a, b]$, i.e. these values need no longer be positive.

We consider the trapezium bounded by $y=\sin x$ and the interval $[\pi, 2\pi]$ of Ox (fig. 91).

One of the primitives of $\sin x$ is $-\cos x$. The difference between its values at $x=2\pi$ and $x=\pi$ is

$$-\cos 2\pi-(-\cos \pi)=-1-1=-2.$$

It will be seen that a negative number has been obtained as a result of using equation (2*). The reasons for this are as follows:

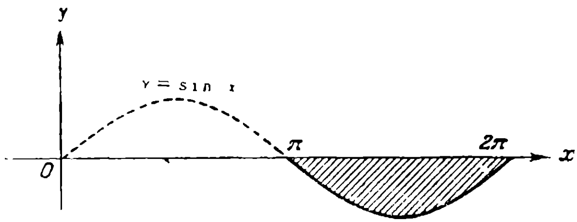


FIG. 91

If the curvilinear trapezium lies below Ox, the ordinates of the curve $y=f(x)$ bounding it have negative values. It may easily be shown that the derivative of the variable area $P(x)$ of such a trapezium is equal to $f(x)$ only when the function $P(x)$ is regarded as negative. It follows from this that the use of expression (2*) for working out the area of a trapezium situated below Ox requires that the area be taken as having a negative value.

We shall agree to take the area of a trapezium lying above Ox as positive, and the area of a trapezium below Ox as negative. In other words, we shall regard an area as an algebraic quantity (as distinct from a geometric quantity).

This condition leads to the fact that cases can occur when the area is zero. Suppose we find, for instance, the area bounded by $y=\sin x$, the ordinates $x=0$, $x=2\pi$, and the segment $[0, 2\pi)$ of Ox.

If we take a primitive of $\sin x$, say $-\cos x$, and subtract its value at $x=0$ from its value at $x=2\pi$, we obtain

$$-\cos 2\pi-(-\cos 0)=-1+1=0.$$

This result is explained by the fact that the trapezium here consists of two symmetrical halves, one above and the other below Ox. The two halves have the same absolute value but opposite signs, so that their algebraic sum is zero (fig. 92).

5. Let the function $f(x)$ be continuous in the interval $[a, b]$

DEFINITION. *The result of subtracting the value of any primitive of $f(x)$ at $x=a$ from its value at $x=b$ is known as the definite integral of $f(x)$ between a and b and is written as*

$$\int_a^b f(x)\,dx.$$

Thus if $F(x)$ is any primitive of $f(x)$, we have by definition

$$\int_a^b f(x)\,dx = F(b) - F(a). \tag{3}$$

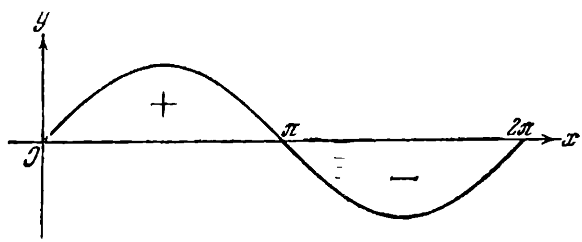


FIG. 92

The difference on the right-hand side of equation (3) is usually denoted by the symbol $[F(x)]_a^b$ or $F(x)|_a^b$, so that we write either

$$\int_a^b f(x)\,dx = [F(x)]_a^b,$$

or

$$\int_a^b f(x)\,dx = F(x)\,|_a^b.$$

For example:

$$\int_0^{\frac{\pi}{2}} \cos x\,dx = [\sin x]_0^{\frac{\pi}{2}} = \sin\frac{\pi}{2} - \sin 0 = 1.$$

6. It follows from the above that the definite integral

$$\int_a^b f(x)\,dx$$

represents from the geometric point of view *the algebraic value of the area of the figure bounded by the curve $y=f(x)$, the straight lines $x=a$, $x=b$ and the intercept $[a, b]$ of the axis of abscissae.*

7. In view of the definition of 5 of the present article evaluation of the definite integral

$$\int_a^b f(x)\,dx$$

is essentially the same as finding any *primitive* of $f(x)$, i.e. finding the indefinite integral $\int f(x)\,dx$. *To evaluate the definite integral*

$$\int_a^b f(x)\,dx,$$

we must first find the primitive $F(x)$ *or the indefinite integral*

$$\int f(x)\,dx=F(x)+C,$$

then evaluate $F(b)-F(a)$, i.e. we apply the formula

$$\int_a^b f(x)\,dx=F(x)\Big|_a^b=F(b)-F(a).$$

Example. Evaluate the definite integral

$$\int_0^1 \sqrt{1-x}\,dx.$$

We first find the indefinite integral

$$\int \sqrt{1-x}\,dx.$$

We do this by substituting $1-x=t$; then $dx=-dt$, and

$$\int\sqrt{1-x}\,dx=-\int t^{\frac{1}{2}}\,dt=-\frac{2}{3}\,t^{\frac{3}{2}}+C=-\frac{2}{3}(1-x)^{\frac{3}{2}}+C.$$

$$\text{Thus } \int_0^1\sqrt{1-x}\,dx=\left[-\frac{2}{3}(1-x)^{\frac{3}{2}}\right]_0^1$$

$$=-\frac{2}{3}\left[(1-1)^{\frac{3}{2}}-(1-0)^{\frac{3}{2}}\right]=-\frac{2}{3}(-1)=\frac{2}{3}.$$

§ 83. Definite integral as the limit of a sum. 1. The problem of finding the area of a curvilinear figure such as we considered in § 82 can be approached in a different manner, by starting from the following considerations:

Let the function $y=f(x)$ be continuous, non-negative and increasing in a segment $[a, b]$ (fig. 93). We consider the figure $aABb$,

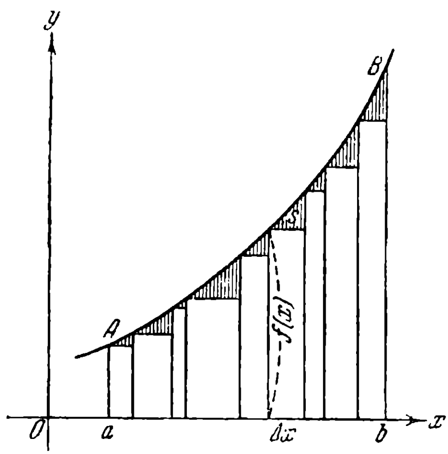


FIG. 93

bounded by an arc of the graph of $y=f(x)$, the straight lines $x=a$, $x=b$ and the segment $[a, b]$ of Ox. The area of this figure will be found by dividing $[a, b]$ into n parts, each of arbitrary length, and then drawing the ordinates to the curve at the points of subdivision. Straight lines are now drawn parallel to Ox from the points where the ordinates meet the curve, so that the whole of the figure is split up into a series of rectangles, each surmounted by a curvilinear triangle (shown shaded in fig. 93). We can find the area of each rectangle, for if we take any one of them and write Δx for its base, its area will be $f(x)\,\Delta x$, where $f(x)$ is its height, this being the ordinate of the point of the curve $y=f(x)$ corresponding to the abscissa of the left-hand end of Δx*.

* We write Δx here both for the actual segment or *sub-interval* and its length.

We have no means of finding the areas of the curvilinear triangles. If s denotes the unknown area of any one of them, we can write the area S of the whole figure as the sum of two terms: the first is the sum

$$\sum_a^b f(x)\,\Delta x$$

of the areas of the rectangles, and the second is the sum $\sum\limits_a^b s$ of the areas of the curvilinear triangles. Thus we now have

$$S \doteq \sum_a^b f(x)\,\Delta x + \sum_a^b s.$$

The Greek letter $\sum$ (sigma) is generally used as an abbreviation to denote a sum. The expression showing what sort of terms are to be summed is placed behind the $\sum$, whilst the letters below and above (a and b) indicate that summation is to be carried out over all the interval from $x=a$ to $x=b$.

If we now let the lengths of all the sub-intervals Δx tend simultaneously to zero (the number n of sub-intervals will obviously increase indefinitely in the process), the second term $\sum\limits_a^b s$ vanishes in the limit, as we shall shortly prove. It is therefore an infinitesimal. The required area S is a constant, whilst the sum $\sum\limits_a^b f(x)\,dx$ is a variable. Hence we have by the definition of a limit:

$$S = \lim_{\Delta x \to 0} \sum_a^b f(x)\,\Delta x.$$

2. We turn to the proof that

$$\lim_{\Delta x \to 0} \sum_a^b s = 0,$$

i.e. we show that the limit of the sum of the areas of the shaded curvilinear triangles is zero when the lengths of all the sub-intervals Δx tend simultaneously to zero.

We displace each triangle parallel to Ox so that its left-hand vertex lies on Oy (fig. 94). The diagram shows that the sum of the areas of all the triangles is less than the area of the horizontally-shaded rectangle. The base of this rectangle is the greatest of the

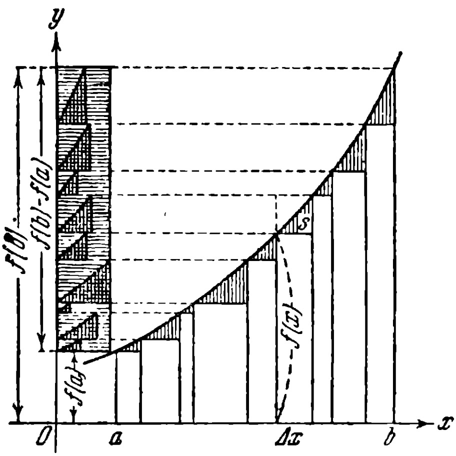


FIG. 94

sub-intervals Δx whilst its height is equal to $f(b)-f(a)$, i.e. the difference between the ordinates of the curve corresponding to the ends b and a of the interval $[a, b]$. Thus

$$0<\sum_{a}^{b} s<[f(b)-f(a)]\,\Delta x, \tag{4}$$

where Δx again denotes the greatest of the sub-intervals into which the total interval $[a, b]$ is divided.

Since we are considering a process in which the lengths of all the sub-intervals tend to zero, the greatest of these must also tend to zero. The difference $f(b)-f(a)$ is a constant. Hence

$$\lim_{\Delta x\to 0}[f(b)-f(a)]\,\Delta x=0.$$

Thus (4) gives

$$\lim_{\Delta x\to 0}\sum_{a}^{b} s=0.$$

Thus *the area S of our figure aABb is equal to the limit of the sum of the areas of the rectangles* $f(x)\,\Delta x$, i.e.

$$S = \lim_{\Delta x \to 0} \sum_a^b f(x)\,\Delta x.$$

3. Now let $y=f(x)$ be a non-negative decreasing function in the interval $[a, b]$ (fig. 95). The area S of $aABb$ is now given as a difference of two sums, viz.

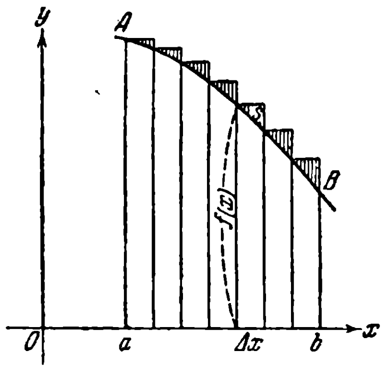


FIG. 95

$$S = \sum_a^b f(x)\,\Delta x - \sum_a^b s,$$

where s denotes the area of a curvilinear triangle.

Using precisely the same arguments as in the previous case, we again find that

$$S = \lim_{\Delta x \to 0} \sum_a^b f(x)\,\Delta x.$$

4. Finally, let $y=f(x)$ both increase and decrease in the interval $[a, b]$ whilst remaining non-negative.

We can now divide $[a, b]$ into sub-intervals in each of which the function is either only increasing or only decreasing. The function $y=f(x)$ show in fig. 96, for instance, is increasing as the abscissa varies from $x=a$ to $x=c$, decreasing from $x=c$ to $x=d$, again increasing from $x=d$ to $x=e$, then finally decreasing again in the

last part of the interval from $x=e$ to $x=b$. We can write on the basis of the above:

$$\text{area } aACc = \lim_{\Delta x\to 0} \sum_a^c f(x)\,\Delta x,$$

$$\text{area } cCDd = \lim_{\Delta x\to 0} \sum_c^d f(x)\,\Delta x,$$

$$\text{area } dDEe = \lim_{\Delta x\to 0} \sum_d^e f(x)\,\Delta x,$$

$$\text{area } eEBb = \lim_{\Delta x\to 0} \sum_e^h f(x)\,\Delta x.$$

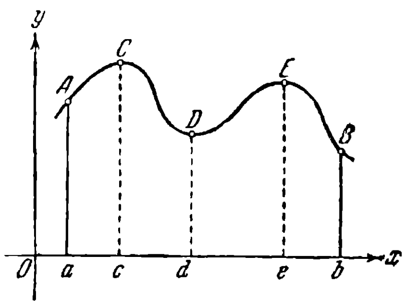


FIG. 96

Obviously, area $aABb$ = area $aACc$ + area $cCDd$ + area $dDEe$ + + area $eEBb$, i.e.

$$\text{area } aABb = \lim_{\Delta x\to 0} \sum_a^c f(x)\,\Delta x + \lim_{\Delta x\to 0} \sum_c^d f(x)\,\Delta x +$$

$$+ \lim_{\Delta x\to 0} \sum_d^e f(x)\,\Delta x + \lim_{\Delta x\to 0} \sum_e^b f(x)\,\Delta x,$$

or, by the theorem regarding the limit of a sum,

$$\text{area } aABb = \lim_{\Delta x\to 0} \{\sum_a^c f(x)\,\Delta x + \sum_c^d f(x)\,\Delta x +$$

$$+ \sum^e f(x)\,\Delta x + \sum_e^b f(x)\,\Delta x\}.$$

But the expression inside the curly brackets is in fact the sum $\sum_a^b f(x)\,\Delta x$. Hence we finally obtain

$$\text{area } aABb = \lim_{\Delta x \to 0} \sum_a^b f(x)\,\Delta x.$$

The result shows that the method we have used for working out the curvilinear area likewise applies when $f(x)$ is either increasing or decreasing over various portions of the interval.

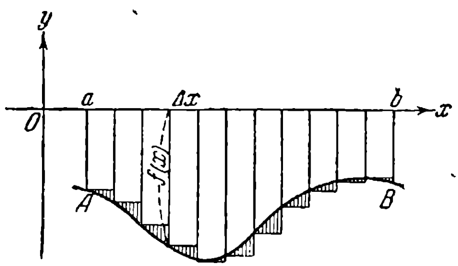


FIG. 97

5. If the curve lies below Ox (fig. 97) the ordinates $y=f(x)$ are negative and the products $f(x)\,\Delta x$ are likewise negative. Hence the sum

$$\sum_a^b f(x)\,\Delta x$$

is negative, and the same can therefore be said of the limit of the sum. We thus find ourselves obliged, as in the previous article, to write a plus or minus sign in front of an area, depending on whether the curvilinear trapezium lies above or below the axis of abscissae.

We established in 6. in § 82 that every definite integral $\int_a^b f(x)\,dx$ can be interpreted as the algebraic value of the area of the curvilinear figure bounded by the curve $y=f(x)$, the straight lines $x=a$, $x=b$, and the intercept $[a\,.\,b]$ of Ox, i.e. precisely the figure the

area of which we have just evaluated as

$$\lim_{\Delta x \to 0} \sum_a^b f(x)\,\Delta x. \tag{4/11}$$

This gives us the extremely important result:

$$\int_a^b f(x)\,dx = \lim_{\Delta x \to 0} \sum_a^b f(x)\,\Delta x; \tag{5}$$

the definite integral $\int_a^b f(x)\,dx$ of the function $f(x)$, continuous in the interval $[a\,,b]$, is the limit of the sum obtained in accordance with the above rule. Turning aside from the special geometric constructions used for forming the sum, we can state the rule for obtaining it as follows: the interval $[a\,,b]$, termed the interval of integration, is arbitrarily divided into n sub-intervals, the value of $f(x)$ at each point of sub-division is multiplied by the length of the corresponding sub-interval Δx, and the resulting products $f(x)\,\Delta x$ are all added together. The sum $S = \sum f(x)\,\Delta x$ thus obtained is known as an *integral sum.*

If the length of each sub-interval Δx tends to zero, the integral sum for a continuous function $f(x)$ tends to a definite limit, as we have already seen, the limit being equal to the definite integral $\int_a^b f(x)\,dx$; whilst it can also be considered from the geometric point of view as the algebraic value of the area of the curvilinear trapezium bounded by the curve $y=f(x)$, the straight lines $x=a$, $x=b$, and the segment $[a\,,b]$ of the axis of abscissae.

This new interpretation of the definite integral is of the greatest impoɪtance in applied mathematics, as we shall see in the following sections.

§ 84. Elementary properties of the definite integral. Three elementary properties follow readily from the definition of the definite integral as the limit of an integral sum.

1. *The integral changes sign when the limits of integration are interchanged.*

This property can be written as

$$\int_a^b f(x)\,dx = -\int_b^a f(x)\,dx.$$

It is based on the following arguments. We have so far considered the case when $a<b$. Now let $b<a$. We now define $\int_a^b f(x)\,dx$ by dividing the interval of integration into sub-intervals Δx and carrying out the process from right to left instead of left to right, so that we must clearly obtain negative Δx. Conversely, division of the interval $[b, a]$ leads to positive values of Δx. Hence the terms $f(x)\,\Delta x$ of the sum

$$\sum_a^b f(x)\,\Delta x$$

will differ only in sign from the corresponding terms of

$$\sum_b^a f(x)\,\Delta x.$$

This means that the sums themselves, and consequently their limits i.e. the integrals

$$\int_a^b f(x)\,dx \quad \text{and} \quad \int_b^a f(x)\,dx,$$

will also differ only in sign.

If $a<b$, we again obtain different signs for the Δx on subdividing $[a, b]$ and $[b, a]$. Hence the relationship

$$\int_a^b f(x)\,dx = -\int_b^a f(x)\,dx$$

till remains valid in this case.

2. *A constant factor can be taken outside the definite integral sign.* This property can be expressed as

$$\int_a^b cf(x)\,dx = c\int_a^b f(x)\,dx.$$

Proof. We have from expression (5) of § 83

$$\int_a^b cf(x)\,dx = \lim_{\Delta x \to 0} \sum_a^b cf(x)\,\Delta x.$$

Since c is a common factor in each term of the sum, we can write

$$\int_a^b cf(x)\,dx = \lim_{\Delta x \to 0} \sum_a^b cf(x)\,\Delta x = \lim_{\Delta x \to 0} c\sum_a^b f(x)\,\Delta x$$

$$= c \lim_{\Delta x \to 0} \sum_a^b f(x)\,\Delta x = c\int_a^b f(x)\,dx.$$

3. *A definite integral of an algebraic sum of functions is equal to the algebraic sum of the definite integrals of the separate functions.*

On confining ourselves to the sum of three terms, we can express this property as

$$\int_a^b [f(x)+\varphi(x)-\psi(x)]\,dx = \int_a^b f(x)\,dx + \int_a^b \varphi(x)\,dx - \int_a^b \psi(x)\,dx.$$

Proof. We have on the basis of expression (5) of § 83:

$$\int_a^b [f(x)+\varphi(x)-\psi(x)]\,dx = \lim_{\Delta x \to 0} \sum_a^b [f(x)+\varphi(x)-\psi(x)]\,\Delta x$$

$$= \lim_{\Delta x \to 0} \sum_a^b f(x)\,\Delta x + \lim_{\Delta x \to 0} \sum_a^b \varphi(x)\,\Delta x - \lim_{\Delta x \to 0} \sum_a^b \psi(x)\,\Delta x$$

$$= \int_a^b f(x)\,dx + \int_a^b \varphi(x)\,dx - \int_a^b \psi(x)\,dx.$$

§ 85. **Principles underlying the use of definite integrals.** A number of problems of an applied nature are solved by the same method as we used in § 83 to find the area of the curvilinear trapezium bounded by the curve $y=f(x)$, the straight lines $x=a$, $x=b$, and the segment $[a, b]$ of the axis of abscissae. We return to this same problem in order to make clear the essential basis of the method.

When we sub-divided the trapezium, we represented its area S as the sum of the areas of rectangles plus the sum of the areas of

curvilinear triangles (fig. 93), viz.

$$S=\sum_a^b f(x)\,\Delta x+\sum_a^b s.$$

Whereas the area of each rectangle, i.e. $f(x)\,\Delta x$, is readily evaluated, we cannot find the area s of a curvilinear triangle, or therefore the sum $\sum_a^b s$ of these areas, by using elementary mathematics. We have shown, however, that $\lim \sum_a^b s=0$ when all the $\Delta x\to 0$. Hence it follows that

$$S=\lim_{\Delta x\to 0}\sum_a^b f(x)\,\Delta x.$$

We thus avoid working out the sum $\sum_a^b s$ of the areas of the curvilinear triangles

We now show that the area s of a curvilinear triangle is an infinitesimal of higher order than Δx (§ 69).

We do this by drawing separately one of the portions into which the trapezium has been sub-divided (fig. 98). As is clear from the

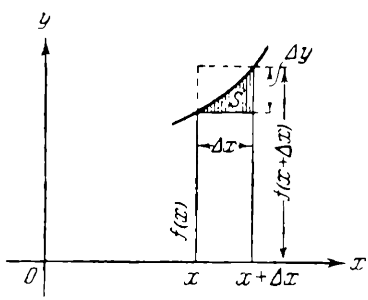


FIG. 98

figure, the shaded area s of the curvilinear triangle is less than the area of the rectangle with sides Δx and Δy, i.e.

$$s<\Delta x\cdot\Delta y.$$

The function $y=f(x)$ is assumed to be continuous as usual, so that $\Delta y \to 0$ as $\Delta x \to 0$. Thus

$$\lim_{\Delta x \to 0} \frac{\Delta x \, \Delta y}{\Delta x} = \lim_{\Delta x \to 0} \Delta y = 0.$$

Hence it will be all the more true that

$$\lim_{\Delta x \to 0} \frac{s}{\Delta x} = 0.$$

which means that s is a higher order infinitesimal than Δx.

We now see, on writing the area S of the curvilinear trapezium as the sum of terms of the form

$$f(x) \, \Delta x + s,$$

that *the quantity s in this expression is an infinitesimal of higher order* than Δx.

The same approach is used in general in problems of the integral calculus. We evaluate a quantity S by expressing it as a sum of terms of the form

$$f(x) \, \Delta x + s,$$

where s cannot be evaluated by elementary methods but is *an infinitesimal of higher order than* Δx (this latter is extremely important), whereas $f(x) \, \Delta x$ can be worked out by elementary mathematics.

As in the case just discussed when s is the area of a curvilinear triangle, it can be shown that the limit of the sum of these higher-order infinitesimals is always zero when all the $\Delta x \to 0$. The quantity S is thus obtained as

$$\lim_{\Delta x \to 0} \sum [f(x) \, \Delta x + s] = \lim_{\Delta x \to 0} \sum f(x) \, \Delta x + \lim_{\Delta x \to 0} \sum s$$
$$= \lim_{\Delta x \to 0} \sum f(x) \, \Delta x.$$

We thus avoid in the general case, as previously, the need for working out the sum $\sum s$.

It follows from this that, having written the quantity S as a sum of terms of the form

$$f(x)\,\Delta x+s,$$

where s is an infinitesimal of higher order than Δx, we can now simply neglect the quantities s and evaluate S at once as

$$\lim_{\Delta x\to 0} \sum f(x)\,\Delta x.$$

This is the principle used in application work.

The independent variable x and the function $y=f(x)$ always have a physical significance in applications of the integral calculus. Thus when considering the motion of a body, x denotes time and y velocity; when working out fluid pressure, as we shall see below, x denotes the depth to which the plate is submerged in the fluid and y the pressure corresponding to the depth x, and so on. But whatever the physical significance of the function, we can always represent it graphically as the ordinate of a curve $y=f(x)$*. Then

$$\lim_{\Delta x\to 0} \sum f(x)\,\Delta x$$

represents the area of some curvilinear trapezium, and since this area is given by the integral

$$\int_a^b f(x)\,dx,$$

the problem of finding the physical quantity S is equivalent to evaluating this definite integral

$$\int_a^b f(x)\,dx.$$

This is the method that we discussed in detail in § 83, when precisely the same proceeding to the limit was used in evaluating the area of a curvilinear trapezium.

§ 86. **Volume of pyramid.** 1. As an example of the use of definite integrals, we shall consider the problem of finding the volume of

* The scale on Ox is here given by the unit of measurement of the physical quantity taken as the independent variable x, whilst the scale on Oy is given by the unit for measuring the physical quantity expressed by the function y.

a pyramid, the formula being first found for the volume of a triangular pyramid in terms of its base area and height.

Let S be the base area and H the height of the pyramid $COAB$ (fig. 99). We divide the height into n arbitrary parts (which may or may not be equal) and draw planes parallel to the base of the pyramid through the points of sub-division. These planes divide

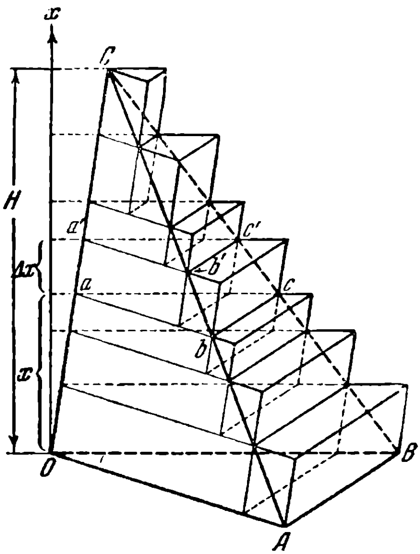


FIG. 99

the total pyramid into n parts. Let the volume of any one of these parts be v. The dividing planes intersect the pyramid in triangles, on the bases of which we construct prisms with their edges parallel to an edge of the pyramid, say the edge OC. There will be two groups of such prisms: one consisting of prisms lying wholly within the pyramid, and the other of prisms projecting out of the pyramid. We shall show that the difference between the volumes of two prisms belonging to different groups but corresponding to the same partial volume v of the pyramid is an infinitesimal of higher order than the common height Δx of the two prisms (see fig. 99).

Let the base of the prism projecting beyond the pyramid be the triangle abc, whilst triangle $a'b'c'$ is the base of the prism lying

inside the pyramid. We write p for the area of triangle abc, and p' for the area of $a'b'c'$. The volume of the projecting prism is

$$p\Delta x = q,$$

whilst the volume of the contained prism is

$$p'\Delta x = q';$$

Hence the difference between the volumes is

$$q - q' = (p - p')\,\Delta x$$

and we therefore have

$$\frac{q - q'}{\Delta x} = p - p'.$$

Clearly $p - p' \to 0$ when $\Delta x \to 0$, and so

$$\lim_{\Delta x \to 0} \frac{q - q'}{\Delta x} = \lim\,(p - p') = 0.$$

which shows that $q - q'$ must be an infinitesimal of higher order than Δx.

But since $q' < v < q$, the difference $q - v$ must be even more an infinitesimal of higher order than Δx. Writing $q - v = \alpha$, we obtain

$$v = q - \alpha,$$

wihere α is an infinitesimal of higher order than Δx. In accordance with the principle explained in the previous section, we neglect the term α and find that

$$V = \lim_{\Delta x \to 0} \sum q = \lim_{\Delta x \to 0} \sum p\,\Delta x = \int_0^H p\,dx.$$

We have by the familiar theorem of elementary geometry:

$$\frac{p}{S} = \frac{(H - x)^2}{H^2},$$

whence

$$p = \frac{(H - x)^2}{H^2} \cdot S,$$

so that

$$V=\int_{0}^{H}\frac{(H-x)^2}{H^2}\cdot S\cdot dx=\frac{S}{H^2}\int_{0}^{H}(H-x)^2\,dx$$

$$=\frac{S}{H^2}\left\{H^2\int_{0}^{H}dx-2H\int_{0}^{H}x\,dx+\int_{0}^{H}x^2\,dx\right.$$

$$=\frac{S}{H^2}\left\{H^3-H^3+\frac{H^3}{3}\right\}=\frac{1}{3}\,S\cdot H.$$

Thus *the volume of a triangular pyramid is equal to a third the base area multiplied by the height.*

2. To find the volume V of a pyramid of height H whose base is an n-sided polygon, we split it up into n triangular pyramids by means of diagonal planes through one edge. These n triangular pyramids are clearly all of the same height H; let their base areas be $s_1, s_2, \ldots, s_n$, and let S be the base area of the original pyramid; since $s_1+s_2+\ \ldots\ +s_n=S$, we have

$$V=\frac{1}{3}\,s_1\cdot H+\frac{1}{3}\,s_2\cdot H+\ldots+\frac{1}{3}\,s_n\cdot H=$$

$$=\frac{1}{3}\,H\,(s_1+s_2+\ldots+s_n)=\frac{1}{3}\,H\cdot S.$$

§ 87. Examples of evaluation of areas. Given the curve $y=f(x)$, it was shown in § 82 that the area S bounded by the curve, by the ordinates of two points of the curve with abscissae a and b, and by the segment $[a\,,\,b]$ of the axis of abscissae is given by

$$S=\int_{a}^{b}f(x)\,dx.$$

We shall consider two examples of evaluation of areas by using this expression.

Example 1. Find the area lying between the parabolas $y^2=2px$ and $x^2=2py$.

Solution. As is clear from fig. 100, the required area will be obtained by subtracting the area of the figure bounded by the parabola $x^2=2py$, the perpendicular to Ox from the point of intersection of the two parabolas, and the Ox axis from the area of the

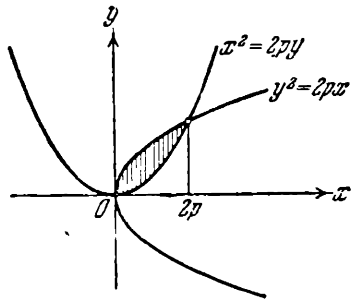


FIG. 100

figure bounded by $y^2=2px$, the same perpendicular, and the Ox axis. These two separate areas will be expressed by two integrals whose limits must be found by working out the abscissae of the points of intersection of the two curves. On solving simultaneously the equations $y^2=2px$ and $x^2=2py$, we obtain $x=0$ and $x=2p$. We thus obtain

$$S=\int_0^{2p}\sqrt{2px}\,dx-\int_0^{2p}\frac{x^2}{2p}\,dx=\sqrt{2p}\int_0^{2p}x^{\frac{1}{2}}\,dx-\frac{1}{2p}\int_0^{2p}x^2\,dx$$

$$=\sqrt{2p}\left[\frac{2}{3}x^{\frac{3}{2}}\right]_0^{2p}-\frac{1}{2p}\left[\frac{x^3}{3}\right]_0^{2p}$$

$$=\sqrt{2p}\cdot\frac{2}{3}\cdot(2p)^{\frac{3}{2}}-\frac{1}{2p}\cdot\frac{8p^3}{3}=\frac{4}{3}p^2.$$

Example 2. Find the area bounded by the ellipse

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1.$$

Solution. Since the area in question is divided into four equal parts by the co-ordinate axes, we can clearly find the required area

by multiplying the area lying in the first quadrant of the axes by four.

The equation of the curve forming one side of the curvilinear trapezium in which we are interested is obtained from the equation of the ellipse as

$$y=f(x)=\frac{b}{a}\sqrt{a^2-x^2}$$

(the + sign is taken in front of the square root because the arc in question lies above Ox (fig. 101). We thus have, writing S for the required area:

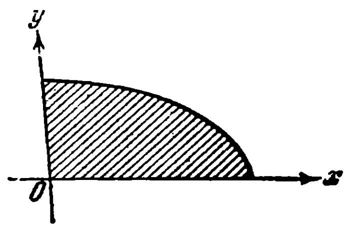


FIG. 101

$$S=4\int_0^a \frac{b}{a}\sqrt{a^2-x^2}\,dx=\frac{4b}{a}\int_0^a\sqrt{a^2-x^2}\,dx.$$

The indefinite integral

$$\int\sqrt{a^2-x^2}\,dx$$

was evaluated in the example on p. 313 as

$$\int\sqrt{a^2-x^2}\,dx=\frac{a^2}{2}\arcsin\frac{x}{a}+\frac{x}{2}\sqrt{a^2-x^2}+C.$$

Hence

$$S=4\,\frac{b}{a}\left[\frac{a^2}{2}\arcsin\frac{x}{a}+\frac{x}{2}\sqrt{a^2-x^2}\right]_0^a$$

$$=4\frac{b}{a}\left(\frac{a^2}{2}\arcsin 1+0-\frac{a^2}{2}\arcsin 0-0\right)$$

$$=4\frac{b}{a}\cdot\frac{a^2}{2}\cdot\frac{\pi}{2}=\pi ab.$$

§ **88. Volume of solid of revolution.** 1. *If the arc AB of a curve rotates about a fixed axis (the axis of revolution), the solid bounded by the surface thus obtained and by the two circular discs formed by the rotation of the perpendiculars from A and B to the fixed axis is known as a solid of revolution* (fig. 102).

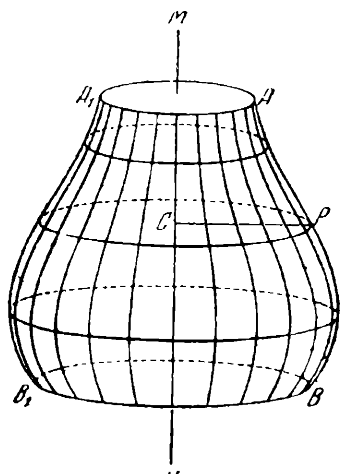


FIG. 102

The arc AB of the rotating curve is known as the *generator* of the surface bounding the solid.

We can regard a solid of this type as being formed by the rotation about the fixed axis of a curvilinear trapezium, bounded by the arc AB of a curve, by the perpendiculars from the ends of the arc to the axis, and by the segment of the axis lying between the feet of these perpendiculars.

Let P be an arbitrary point on the generator, and let PC be the perpendicular from P to the axis of revolution. Obviously, neither

the length of PC nor the angle MCP nor the position of the point C vary during the rotation. Thus each point of the generator describes on rotation a circle lying in a plane perpendicular to the axis of revolution.

2. The simplest solid of revolution is the right circular cylinder. We shall deduce the formula for the volume V of the cylinder in terms of its radius r and height H.

We inscribe a regular n-sided prism in the cylinder and circumscribe a similar regular prism with the same number of sides about the cylinder. Let s denote the base area of the inscribed, and S the base area of the circumscribed prism. Obviously, the prisms have the same height H as the cylinder. We know from elementary geometry that the volumes of the inscribed and circumscribed prisms are respectively sH and SH. We form the difference

$$SH-sH=(S-s)\,H.$$

We also know from geometry that, on indefinite increase of the number of sides of the base polygons of the prisms, and on indefinite approach of the length of side to zero, the difference $S-s\to 0$.

Let V be the volume of the cylinder. Since evidently $V<SH$, we have

$$0<V-sH<SH-sH=(S-s)H.$$

Hence it follows that, as $S-s\to 0$, it will be even more true that $V-sH\to 0$. Since V is constant, whilst $V-sH$ is an infinitesimal, we have

$$V=\lim sH=H\cdot\lim s.$$

But we know (likewise from elementary geometry) that $\lim s=\pi r^2$, so that we finally obtain

$$V=\pi r^2\,H.$$

3. We now deduce the formula for a solid of revolution, given the equation of its generator.

We take the arc of the curve $y=f(x)$ lying between the points whose abscissae are a and b $(a<b)$. We want to find the volume of the solid obtained by rotation of this arc about Ox.

We use the same approach as in finding the area of a plane figure, and imagine the whole solid split up into a large number of portions by planes perpendicular to Ox. The volume of each portion will now consist of the volume of the cylinder formed by rotation of the rectangle of base Δx and height y (fig. 103) plus the volume obtained by rotation of the curvilinear triangle situated above the rectangle (the triangle is shown shaded in fig. 103). It can be shown that the volume obtained by revolution of the curvilinear triangle

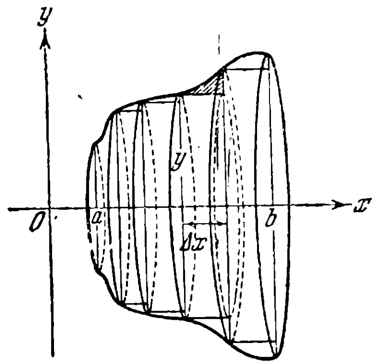


FIG. 103

is an infinitesimal of higher order than Δx. We neglect all the volumes of this latter type in accordance with the principle stated in § 85 and express the volume V of the solid of revolution as the limit of the sum of the volumes of the cylinders. Since the volume of a cylinder of radius y and height Δx is $\pi y^2 \Delta x$, we have

$$V = \lim_{\Delta x \to 0} \sum_a^b \pi y^2 \Delta x = \pi \int_a^b y^2 \, dx = \pi \int_a^b [f(x)]^2 \, dx. \tag{6}$$

Example. Find the volume of the solid obtained by revolution of the ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ about Ox (fig. 104).

Solution. Since $y^2 = \frac{b^2}{a^2}(a^2 - x^2)$, we obtain by equation (6):

$$V = \pi \frac{b^2}{a^2} \int_{-a}^{+a} (a^2 - x^2) \, dx = \pi \frac{b^2}{a^2} \left[a^2 x - \frac{x^3}{3} \right]_{-a}^{+a} = \frac{4}{3} \pi a b^2.$$

§ 89. Volumes of cone, truncated cone, sphere and segment of sphere. 1. *Volume of a right circular cone. The right circular cone can be regarded as the solid obtained by revolution of a right-angled*

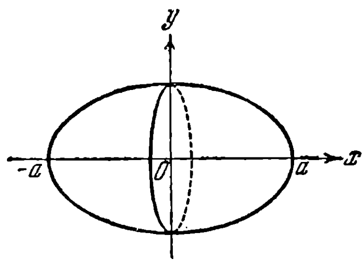


FIG. 104

triangle OAB about an axis coinciding with one of the adjacent sides, say OB.

We deduce the expression for the volume of the cone in terms of its height H and the radius of its base r.

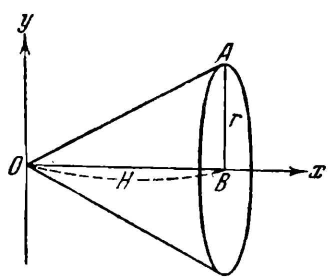


FIG. 105

We take the vertex O of the cone as origin and direct Ox along the axis of the cone (fig. 105). The generator OA of the cone is a straight line whose slope is clearly $\frac{r}{H}$. Thus the equation of OA will be

$$y = \frac{r}{H} x .$$

Writing V for the volume of the cone, we have by equation (6) of § 88:

$$V=\pi\int_0^H y^2\,dx=\pi\int_0^H \frac{r^2}{H^2}\,x^2\,dx=\frac{\pi r^2}{H^2}\int_0^H x^2\,dx=\frac{\pi r^2}{H^2}\cdot\left[\frac{x^3}{3}\right]_0^H=\frac{1}{3}\,\pi r^2 H.$$

2. *Volume of truncated cone.* We shall regard the truncated cone as the solid of revolution formed by rotating the trapezium $OABB_1$ about the side OB_1 perpendicular to the bases OA and BB_1 (fig. 106).

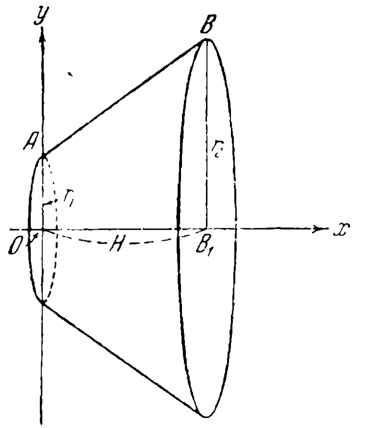


FIG. 106

We deduce the expression for the volume V of the truncated cone in terms of the radii $r_1=OA$, $r_2=B_1B$ of its bases and its height $H=OB_1$.

The slope of the generator AB is easily seen to be $\frac{r_2-r_1}{H}$.

Hence the equation of the generator is

$$y=\frac{r_2-r_1}{H}\,x+r_1.$$

We now have by equation (6) of § 88:

$$V=\pi\int_0^H\left(\frac{r_2-r_1}{H}\,x+r_1\right)^2 dx=\pi\left\{\frac{(r_2-r_1)^2}{H^2}\int_0^H x^2\,dx+\right.$$

$$+2r_1\frac{r_2-r_1}{H}\int_0^H x\,dx+r_1^2\int_0^H dx\Big\}=\pi\left\{\frac{(r_2-r_1)^2}{H^2}\cdot\frac{H^3}{3}+\right.$$

$$\left.+2r_1\frac{r_2-r_1}{H}\cdot\frac{H^2}{2}+r_1^2H\right\}=\frac{1}{3}\pi r_2^2H+\frac{1}{3}\pi r_1^2H+\frac{1}{3}\pi r_1r_2H.$$

Thus the volume of a truncated cone is the sum of the volumes of three cones with the same height as the truncated cone and bases equal respectively to the major and minor bases of the truncated cone and to the circle whose area is the geometric mean of the areas of the major and minor bases.

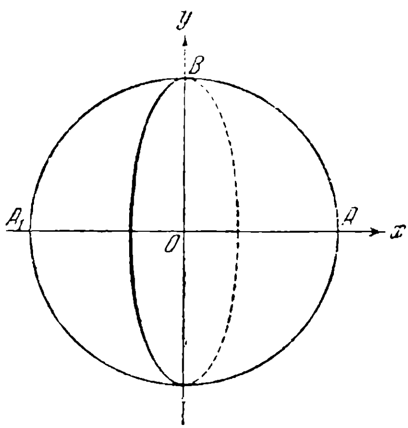


FIG. 107

3. *Volume of sphere.* We shall regard the sphere as the solid obtained when a semi-circle revolves about its diameter, i.e. when the arc ABA_1 of the circle

$$x^2+y^2=r^2$$

revolves about Ox (fig. 107). We have from the equation of the circle, $y^2=r^2-x^2$. Writing V for the volume of the sphere, we have by (6) of § 88:

$$V=\pi\int_{-r}^{+r}(r^2-x^2)\,dx=\pi\left\{r^2\int_{-r}^{+r}dx-\int_{-r}^{+r}x^2\,dx\right\}$$

$$=\pi\left\{r^2[x]_{-r}^{+r}-\left[\frac{x^3}{3}\right]_{-r}^{+r}\right\}=\pi\left(2r^3-\frac{2}{3}r^3\right)=\frac{4}{3}\pi r^3.$$

4. *Volume of segment of sphere.* The segment of a sphere can evidently be looked on as the solid obtained by revolution of the segment BAB' of a circle about the diameter A_1A of the circle (fig. 108). The piece CA of the diameter of the sphere is termed

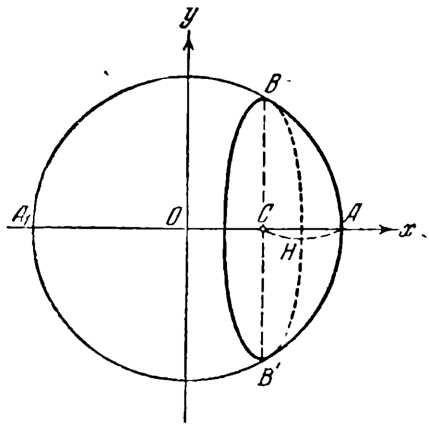


FIG. 108

the *height* of the segment. The plane that cuts off the segment from the sphere intersects the sphere in a circle known as the *base* of the segment.

We shall obtain the formula for the volume V of the segment in terms of the radius r of the sphere and the height $CA=H$ of the segment.

The equation of the circle of which BAB' is a segment is

$$x^2+y^2=r^2.$$

Hence $y^2=r^2-x^2$.

We now have for V, by (6) of § 88,

$$V=\pi\int_{r-H}^{r}(r^2-x^2)\,dx=\pi\left\{r^2[x]_{r-H}^{r}-\left[\frac{x^3}{3}\right]_{r-H}^{r}\right\}$$

$$=\pi\left\{r^2[r-(r-H)]-\frac{1}{3}[r^3-(r-H)^3]\right\}$$

$$=\pi\left\{r^2H-\frac{1}{3}(r^3-r^3+3r^2H-3rH^2+H^3)\right\}$$

$$=\pi\left(r^2H-r^2H+rH^2-\frac{1}{3}H^3\right)=\pi H^2\left(r-\frac{1}{3}H\right).$$

Thus *the volume of a segment of a sphere is equal to the volume of a cylinder whose base radius is equal to the height of the segment and whose height is equal to the radius of the sphere minus a third the height of the segment.*

§ 90. **Work done by a force.** If a body moves a distance x along a straight line under the action of a force F whose direction is the same as the direction of motion of the body, the work done by the force is defined as the product $F.x$.

If the force is variable, the work done can only be found by proceeding to some limit.

Let the body move along the Ox axis from the point A $(x=a)$ to the point B $(x=b)$ (fig. 109) under the action of a variable force F, directed along Ox, F being a function of x, i.e. $F=f(x)$.

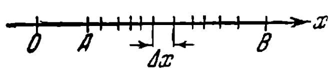


FIG. 109

We find the work P, done by the force F when the body moves over the distance AB, by proceeding in a manner similar to that adopted when we considered the geometrical applications of integrals.

We arbitrarily divide the total distance AB into n parts (Δx) and write Δx for the length of any one of these parts. We now consider approximately that the body moves over each part (Δx) under the action of a constant force, equal to the value of $F=f(x)$ at say the left-hand end of (Δx). The work done by the force over

each (Δx) is now given by $f(x)\,\Delta x$ (called the elementary work). This approximate value for the work differs from the strict value by an infinitesimal of higher order than Δx. In accordance with the general rule stated in § 85, this error can be neglected, so that the work P done by the force F over the whole distance AB is given by the limit of the sum of the elementary works on indefinite increase of the number n of parts (Δx) (assuming that all the Δx tend to zero). We thus obtain

$$P=\lim_{\Delta x\to 0}\sum_a^b f(x)\,\Delta x=\int_a^b f(x)\,dx.$$

Example. The compression of a spiral spring is proportional to the force applied. Find the work done on compressing the spring 3 cm, given that a force of 1 kg is required for a compression of 0·5 cm.

Solution. Let the compression be x, measured in metres. Then the force required for this compression will be kx, where k is a coefficient of proportionality. With $x=0{\cdot}005$ m, the force $f(x)$ $=f(0{\cdot}005)=k\,0{\cdot}005=1$, whence $k=\dfrac{1}{0{\cdot}005}=200$ and

$$f(x)=200x.$$

By the above formula, the work done by the force on compressing the spring 3 cm ($=0{\cdot}03$ m) is

$$\int_0^{0.03} f(x)\,dx=\int_0^{0.03} 200x\,dx=100\,[x^2]_0^{0.03}=0.09\text{ kg}.$$

§ **91. Pressure of fluid.** The fluid pressure on a horizontal plate submerged at a depth h below the free surface of the fluid is equal to the weight of the column of fluid supported by the plate. Thus, if S denotes the area of the plate, γ is the specific gravity of the fluid, and P is the pressure, we have

$$P=\gamma\cdot S\cdot h.$$

Starting from the fact that the fluid pressure is the same in all directions, we shall find the pressure exerted on a vertically sub-

merged plate by following the same method as used for solving the problems of the previous sections.

We imagine the plate split up by straight lines parallel to the free surface of the fluid into a large number n of narrow strips (fig. 110) and assume that the pressure is the same at every point

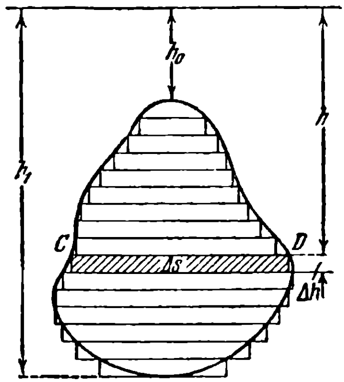


FIG. 110

of any given strip. This assumption, which is obviously not strictly correct, since the pressure will be greater at the bottom than at the top of a strip, is more justified if the strip is fairly narrow. We shall take the strips to be rectangular for the sake of simplicity. Let the area of any one of them be Δs. It can be shown that the error involved in these assumptions is an infinitesimal of higher order than Δs, so that we can neglect it in accordance with our general rule.

On writing γ for the specific gravity of the fluid, and taking into account the above assumptions, we find that the pressure on a strip at a depth h is given approximately by $\gamma h \Delta s$.

Moreover, we have from fig. 110:

$$\Delta s = CD\,\Delta h\ ;\quad \gamma h\,\Delta s = \gamma h \cdot CD \cdot \Delta h. \tag{6/14}$$

On increasing indefinitely the number of strips n and passing to the limit as $n \to \infty$ (on the assumption that all the heights of the strips tend to zero), we find for the pressure P on the plate:

$$P = \lim \sum_{h_0}^{h_1} \gamma h CD\, \Delta h = \int_{h_0}^{h_1} \gamma h CD\, dh, \tag{7}$$

where h_0 is the depth of the uppermost, and h_1 the depth of the lowest point of the plate. To evaluate the integral obtained, we must be able to express the length CD of an arbitrary strip in terms of h, i.e. we have to know the shape of the plate.

Example. Find the pressure exerted on a dam of trapezium shape with parallel bases of 40 m and 15 m respectively. The height of the dam is 8 m and its upper base lies on the surface of the water. (We shall assume below that the weight of a cubic metre of water is 1000 kg = 1 t.

Solution. We have by formula (7), $P = 1000 \int_0^8 hCD\, dh$.

Here, from fig. 111,

$$CD = CF + FD = CF + 15.$$

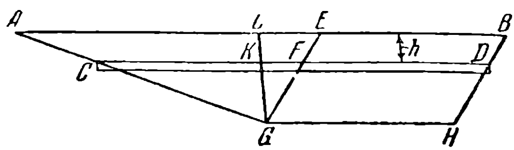


FIG. 111

We obtain from the similar triangles AGE and CGF:

$$\frac{CF}{AE} = \frac{GK}{GL} \quad \text{or} \quad \frac{CF}{25} = \frac{8-h}{8}, \quad \text{whence} \quad CF = \frac{25(8-h)}{8}.$$

Hence

$$CD = \frac{25(8-h)}{8} + 15 = \frac{320 - 25h}{8} = 40 - \frac{25}{8} h$$

and

$$P = 1000 \int_0^8 \left(40 - \frac{25}{8} h\right) h\, dh = 746 \frac{2}{3} t\,.$$

EXERCISES

On §§ 82, 84, 87.

Evaluate the following definite integrals:

1. $\int_2^6 x\,dx.$ Ans. 16.

2. $\int_0^2 3x^4\,dx.$ Ans. 19.2.

3. $\int_{-1}^3 (1-2x+3x^2)\,dx.$ Ans. 24.

4. $\int_{-a}^{+a} x^4\,dx.$ Ans. $\frac{2a^5}{5}.$

5. $\int_{-a}^{+a} x^3\,dx.$ Ans. 0.

6. $\int_{-b}^{+b} (x^2-1)\,dx.$ Ans. $\frac{2b}{3}(b^2-3).$

7. $\int_1^2 \frac{dx}{x^2}.$ Ans. $\frac{1}{2}.$

8. $\int_2^3 \left(\sqrt{x}+\frac{1}{\sqrt{x}}\right) dx.$ Ans. $\frac{2}{3}(6\sqrt{3}-5\sqrt{2}).$

9. $\int_1^2 \frac{2x^2+1}{x}\,dx.$ Ans. $3+\ln 2.$

10. $\int_0^1 \frac{dx}{\sqrt{1-x^2}}.$ Ans. $\frac{\pi}{2}.$

11. $\int_0^1 \frac{dx}{1+x^2}.$ Ans. $\frac{\pi}{4}.$

12. $\displaystyle\int_0^2 \frac{dx}{4+x^2}$. Ans. $\dfrac{\pi}{8}$.

13. $\displaystyle\int_{\frac{\pi}{6}}^{\frac{\pi}{4}} \cos 2x\,dx$. Ans. $\dfrac{2-\sqrt{3}}{4}$.

14. $\displaystyle\int_{\frac{\pi}{4}}^{\frac{\pi}{3}} \tan^2 x\,dx$. Ans. $\sqrt{3}-1-\dfrac{\pi}{12}$.

15. $\displaystyle\int_0^{\frac{\pi}{2}} \cos^2 x\,dx$. Ans. $\dfrac{\pi}{4}$.

16. $\displaystyle\int_0^{\frac{\pi}{2}} \sin^3 x\,dx$. Ans. $\dfrac{2}{3}$.

17. $\displaystyle\int_0^{\frac{\pi}{2}} \sin^3 x \cos^2 x\,dx$. Ans. $\dfrac{4}{5}$.

18. $\displaystyle\int_0^{\frac{\pi}{2}} \sin^2 x \cos^2 x\,dx$. Ans. $\dfrac{\pi}{16}$.

19. $\displaystyle\int_0^1 \frac{x^2\,dx}{1+x^3}$. Ans. $\dfrac{1}{3}\ln 2$.

20. $\displaystyle\int_0^1 \frac{x^2\,dx}{1+x^6}$. Ans. $\dfrac{\pi}{12}$.

21. $\displaystyle\int_1^e \frac{\ln x\,dx}{x}$. Ans. $\dfrac{1}{2}$.

Find the areas of the figures whose boundaries are given by the following equations:

22. $y=5x$, $x=2$, $y=0$. Ans. 10.

23. $y=3x-1$, $x=2$, $x=4$, $y=0$. Ans. 16.

24. $x-y+1=0$, $3x+2y-12=0$, $y=0$. Ans. 7.5.

25. $x-4y+2=0$, $x+y-3=0$, $y=0$. Ans. 2.5.

26. $y^2=4x$, $x=4$, $x=9$. Ans. $25\frac{1}{3}$.

27. $y=x^3$, $y=2x$. Ans. 2.

28. $xy=a^2$, $x=a$, $x=2a$, $y=0$. Ans. $a^2 \ln 2$.

29. $y^2=ax$, $x^2=by$. Ans. $\frac{ab}{3}$.

30. $y^2=4{\cdot}5x$, $3x-4y=0$. Ans. 8.

31. $y^2=8x$, $2x-3y+8=0$. Ans. $\frac{4}{3}$.

32. $x^2=9y$, $x-3y+6=0$. Ans. $13\frac{1}{2}$.

33. $4x^2-8x-y+5=0$, $2x-y+1=0$. Ans. $2\frac{1}{4}$.

34. $x^2-6x-4y+13=0$, $x-2y-1=0$. Ans. $\frac{1}{3}$.

35. $3y^2-16x+32=0$, $4x-3y-8=0$. Ans. 2.

36. $6y^2-25x-50=0,$

$5x-6y+10=0.$ Ans. 5.

37. $4x^2-9y+18=0,$

$2x^2-9y+36=0.$ Ans. 8.

38. $5x^2-60x+4y+160=0,$

$x^2-12x+2y+32=0.$ Ans. 8.

On § 88.

39. Find the volume of the paraboloid of revolution formed by revolving about Ox the arc of the parabola $y^2=4x$ lying between the points $(0, 0)$ and $(4, 4)$.

Ans. 32π.

40. Find the volume of the solid formed by revolving about Ox the arc cut off from the parabola $y=x^2-4$ by Ox.

Ans. $\frac{512}{15}\pi$.

41. Find by integration the volume of the solid formed by revolving about Ox the intercept of the straight line $4x-5y+3=0$ cut off by the co-ordinate axes.

Ans. $0{\cdot}09\,\pi$.

42. Find the volume of the solid formed by revolving about its minor axis the ellipse $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$ $(a>b)$.

Ans. $\frac{4}{3}\pi a^2 b$.

Hint. The volume V of a solid formed by revolving about Oy an arc of the curve $y=f(x)$ is given by

$$V=\pi\int_a^b x^2\,dy.$$

43. Find the volume of the solid formed by revolving about its transverse axis the arc of the hyperbola $\frac{x^2}{a^2}-\frac{y^2}{b^2}=1$, the abscissae of the ends of the arc being $x=a$ and $x=c$ $(c>a)$.

Ans. $\frac{\pi b^2}{3}\left(\frac{c^3-3a^2c}{a^2}+2a\right)$.

44. Find the volume of the solid formed by revolving about Ox the arc of $y=\sin x$ lying between the origin and $(\pi, 0)$.

Ans. $\frac{1}{2}\pi^2$.

On § 90.

45. The force required to stretch a metal rod of length a to a length $a+x$ is $\frac{xk}{a}$, where k is a constant. Find the work done in stretching the rod from a length a to a length b.

Ans. $\frac{k(b-a)^2}{2a}$.

46. A gas is enclosed in a cylinder with a moving piston, the area of which is equal to A. Using Boyle's law: $pv=k$, find the work done by the gas pressure when the gas volume increases from v_1 to v_2.

Ans. $k\ln\frac{v_2}{v_1}$.

Hint. The force exerted by the gas on the piston is pA, where p is the pressure per unit area. When the piston is pushed out a distance dx, the elementary work done is $pAdx$. But Adx is the increment dv in the gas volume. Hence $pAdx=\frac{k}{v}dv$, and the work done is given by $\int_{v_1}^{v}\frac{k}{v}dv$.

47. By Newton's law, the force of attraction is inversely proportional to the square of the distance. A material particle in a

state of rest attracts a second particle which moves along a straight line from a distance r_1 to a distance r_2 from the first particle. Find the work done by the force of attraction.

Ans. $\mu\left(\frac{1}{r_2}-\frac{1}{r_1}\right)$, where μ is a constant of proportionality.

On § 91.

48. Find the force exerted on a rectangular plate of base 8 cm and height 12 cm vertically submerged in water so that the upper base of the plate is 5 cm below the free surfase of the water.

Ans. 1056 g.

49. Find the force exerted on a triangular plate of base 10 cm and height 4 cm submerged vertically in water so that its vertex lies in the surface of the water.

Ans. $53\frac{1}{3}\,g$.

50. Find the force exerted on a triangular plate of base 8 cm and height 6 cm vertically submerged in water so that its vertex lies above the base and at a distance of 3 cm below the free surface of the water.

Ans. 163 g.

51. A plate consisting of a parabolic segment of base 15 cm and height 3 cm is vertically submerged in water so that the vertex is on the surface of the water. Find the force exerted on the plate.

Ans. 54 g.

CONCLUSION

THE fact that mathematics has achieved such enormous importance, as a tool for research into the external world and as a means of solving practical scientific and engineering problems is due to the creation in the second half of the 17th century of an apparatus—analitic geometry and infinitesimal calculus—whereby *variable* quantities could be studied.

Variables make their first appearance in these pages as the current co-ordinates of a point on a plane (see Ch. 2, § 5). From

current co-ordinates we passed on to the representation of various curves by means of equations, and these enabled us to investigate the properties of geometric curves, firstly by means of algebra and trigonometry, then later by the methods of mathematical analysis (differential calculus).

Geometry and analysis, which would appear to be quite distinct subjects, are thus shown to be intimately related, thanks to the introduction of variables into mathematics. The relationship has been of mutual benefit to both subjects by contributing to their more intensive development.

We now know that the equation of a curve expresses a functional relationship between two variables—the abscissa and ordinate of a point of the curve. The function, or functional relationship between certain variables, reflects an interconnexion between real physical entities, and it becomes possible to investigate physical processes mathematically. We have encountered a number of examples of functions representing physical processes in the course of these pages: the law of motion of a material particle, the speed of a chemical reaction, the law for electrical current variation, and so on.

The differential and integral calculus owe their development to the fact that a study of natural phenomena and of engineering processes leads in the vast majority of cases to mathematical problems of *two types,* one type being solved essentially by the methods of the *differential* calculus and the other by the methods of the *integral* calculus.

We are already familiar with the fact that the differential calculus is used for finding the *rate of change* of a process: the velocity of a moving particle, the velocity of a chemical reaction, etc. Such problems arise when we know the course of the process as a whole (e.g. the law of the chemical reaction) and we want to know the velocity at which the process is taking place at any given instant. Problems of differentiation can be stated in general terms as those of determining the instantaneous, localised character of a process (or to express the matter geometrically, the nature of the process "at a point") on the basis of its known properties "as a whole" (i.e. over a given distance, interval of time, etc.).

A preliminary historical note is needed before we can explain the principles underlying the use of the integral calculus.

For purely pedagogic reasons, the systematic exposition of the integral calculus to be found in the present book does not entirely accord with the historical development of the subject. Not only is the definite integral of earlier origin than the indefinite, but it was known first as the limit of an integral sum (see § 83) and not as the difference between the values of a primitive (see § 82). The explanation lies in the fact that a proceeding to the limit of this kind represents the most natural approach to solving a very wide range of practical problems, as for instance that of finding the work done by a force (see § 90). This particular problem may be described as follows: we know the value of the force as a function of the abscissa of a point which moves along a straight line under the action of the force, and we want to find the work done by the force when the point moves over a given distance. This represents a typical problem of the integral calculus. Starting from this particular example, we can describe a problem of the integral calculus in general terms as one in which we want to find the properties of some process "as a whole" (over a given distance, interval of time, etc.) on the basis of its characteristic features at a point (its instantaneous, localised features).

We already know that finding the work done by a force reduces to finding the limit of an integral sum, i.e. to evaluating a definite integral (§ 90). The same can be said of every problem of the type in question.

Problems of this type were always solved by direct evaluation of the limit of the integral sum in the first stage of development of the integral calculus; a number of problems had in fact been solved in classical times by this method (Archimedes). Such an approach means, however, that each time a new integration problem arises special methods have to be worked out for finding the limit of the particular integral sum. There was no general method in the early days for finding the limit of an integral sum.

A simple comparison of problems of differentiation and integration shows that the two types of problem are the *inverse of each other* in the full sense of the word. This fact did not lead, however,

to the realization of the essential connexion between the differential and integral calculus until the works of Newton and Leibniz appeared. The connexion lies in the fact that the derivative of a variable area, i.e. the derivative of the integral $\int_a^x f(x)\,dx$ with respect to x is equal to the integrand $f(x)$. After this, it was realized that the definite integral $\int_a^b f(x)\,dx$ is not only the limit of an integral sum but is also equal to the difference between the values of any primitive $F(x)$ of the integrand $f(x)$ for values of the argument equal to the upper and lower limits of integration, i.e.

$$\int_a^b f(x)\,dx = F(b) - F(a), \tag{1}$$

where $F(x)$ is a function such that $F'(x) = f(x)$.

Thanks to this discovery, it was possible to replace direct evaluation of a definite integral as the limit of a sum by an operation the inverse of differentiation. A general method could thus be developed for solving problems of both types.

This is why we began our study of the integral calculus by considering the operation the inverse of differentiation, i.e. we first considered indefinite integrals.

The great merit of Newton and Leibniz lies in the fact that they fully appreciated the intimate connexion between the differential and integral calculus and then made use of it to produce a unified mathematical theory—the analysis of infinitesimals (see p. 13). The above expression (1) is often referred to in mathematical literature as the Newton-Leibniz formula, a name which acts as a permanent reminder of the achievements of these great men.

The connexion between the differential and integral calculus serves as an example of the dialectical development of the sciences: two subjects that appear to be opposed in the early stages of their development later combine into a dialectic unity.

These closing remarks have been made with the aim of underlining some of the main points in our treatment of mathematical theory.

ADDENDUM

CHAPTER 11

DIFFERENTIAL EQUATIONS OF THE FIRST ORDER WITH SEPARABLE VARIABLES

§ 92. **Definitions.** *Let x be the independent variable and y an unknown function of x; an equation connecting x and y and derivatives (or differentials) of any order of y is termed a differential equation.* For instance,

$$y''+y=0 \tag{1}$$

is a differential equation (this does not contain the argument x explicitly).

Every function satisfying a differential equation is a *solution* of it. For example, $\sin x$ is a solution of equation (1). For we obtain, on substituting $\sin x$ for y in the left-hand side of (1),

$$(\sin x)''+\sin x=-\sin x+\sin x,$$

i.e. we get zero on the left-hand side, so that the equation reduces to the numerical identity $0=0$.

The *order* of a differential equation is the same as the highest order of any derivative appearing in the equation. Thus

$$xy'+y=x\sin x$$

is a differential equation of the *first order.* Equation (1) is of the *second order,* since the highest order derivative appearing in it is of the second order.

It may easily be seen that the function

$$y=C_1\sin x+C_2\cos x, \tag{2}$$

where C_1 and C_2 are arbitrary constants, is also a solution of

equation (1).* A solution containing two arbitrary constants of a second order differential equation is termed a *general* solution; thus (2) represents a general solution of equation (1).

Every solution which is obtained by assigning definite numerical values to the arbitrary constants in the general solution of a second order differential equation is termed a *particular* solution of the equation. Thus we get the particular solution $y=\sin x$ by setting $C_1=1$ and $C_2=0$ in the general solution (2) of equation (1).

The general and particular solutions of equations of the 1st, 3rd, 4th and higher orders are defined in the same way.

Apart from particular solutions, a differential equation may have solutions which cannot be obtained no matter what values are assigned to the arbitrary constants of the general solution; such solutions are described as *singular*.

To *solve* or *integrate* a differential equation means to find all its solutions.

There is a wide range of types of differential equation. We shall confine our attention to the most elementary type of first order equation, with *separable variables*, and shall indicate the method of finding the general solution (§ 94). To start with, we shall examine some problems leading to differential equations.

§ 93. Examples of problems leading to differential equations.
1. Find the equation of a curve such that the slope of its tangent is always equal to twice the abscissa of the point of contact.

Solution. From the data given we find the equation

$$\frac{dy}{dx}=2x.$$

On multiplying both sides by dx, we obtain

$$dy=2x\,dx.$$

* To be precise, $y=C_1\sin x+C_2\cos x$ defines an infinite set of functions, and since the arbitrary constants C_1, C_2 can be assigned any numerical values the relationship is said to define a *family* of functions satisfying equation (1).

Assuming that y is a function satisfying this equation, we can integrate both sides to give

$$\int dy = \int 2x\, dx:$$

whence we find

$$y = x^2 + C,$$

where C is an integration constant.

The function obtained is the general solution of the differential equation, containing the arbitrary constant C. Instead of finding a single function, we have in fact found an infinite set of functions, differing from each other in the value of C, i.e. we have a family of functions (see footnote on p. 369).

In geometrical terms, instead of a single curve we have obtained a family of curves, each corresponding to a different value of C (fig. 112).

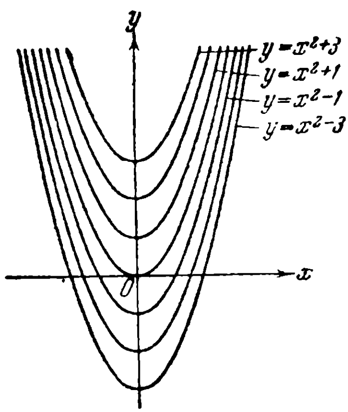


FIG. 112

If it were stated in the problem that the curve must pass through a given point of the plane, say the point (2 , 3), we should obtain a unique solution, since only one parabola of the family $y = x^2 + C$ will satisfy the condition imposed. For the equation of the required curve must be satisfied by the co-ordinates (2 , 3), and this can only happen for a unique value of C, found by substituting (2 , 3)

in the equation of the family $y=x^2+C$. We obtain $3=4+C$, whence $C=-1$, and the equation of the required curve is

$$y=x^2-1.$$

We have found C *in accordance with the initial values* (see § 79).

2. A particle moving from rest has a velocity $v=gt$ at the instant t, where g is a constant. Find the law of motion, i.e. the relationship between the distance s traversed by the particle and time t.

Solution. We know (§ 41) that the velocity v is the derivative of the path s with respect to time t, i.e. $v=\frac{ds}{dt}$. Since we know from the problem that $v=gt$, we arrive at the differential equation

$$\frac{ds}{dt}=gt\,.$$

We rewrite this as $ds=gtdt$ and integrate to give

$$\int ds=\int gt\,dt\,,$$

whence we obtain $s=\frac{1}{2}gt^2+C$.

We find the value of the constant C by observing that $s=0$ at $t=0$. Substitution of these values in the above equation gives us $C=0$.

The required law of motion is thus given by

$$s=\frac{1}{2}\,gt^2\,.$$

3. The rate of cooling of a body in air is proportional to the temperature difference between the body and the air. The air temperature is 20°C. We are told that the body cools from 100° to 60° in 20 min. How long does it take for the body to cool to 30°?

Solution. Let the body temperature be T. The rate of change of T with time t is equal to the derivative $\frac{dT}{dt}$ (§ 41). Since the temperature T is decreasing, $\frac{dT}{dt}$ will be the derivative of a decreas-

ing function and therefore negative (see § 61). But since the problem is concerned with the rate of temperature change when the body is cooling, it follows that we have to regard the velocity of this physical process as a positive quantity, i.e. as $\frac{dT}{dt}$ taken with reversed sign. The rate of cooling of the body is thus given by $-\frac{dT}{dt}$.

On the other hand, the law mentioned in the problem tells us that the rate of cooling is given by

$$k\,(T-20),$$

where k is a coefficient of proportionality ($k>0$).

Comparison of the two expressions for the rate of cooling leads us to the differential equation

$$-\frac{dT}{dt}=k\,(T-20)\,.$$

We rewrite this as

$$\frac{dT}{T-20}=-k\,dt$$

and integrate to obtain $\ln\,(T-20)=-kt+C$.

By the definition of logarithm, this last equation can be written as

$$T-20=e^{-kt+C},$$

or

$$T-20=e^{-kt}\cdot e^{C}\,.$$

On writing $e^{C}=c$, we obtain

$$T-20=ce^{-kt}\,.$$

The constants c and k are still unknown in this solution. We find c by observing that the body temperature T is 100° at $t=0$. We obtain on substituting these values $80=ce^{\circ}$, whence $c=80$.

We thus have

$$T-20=80e^{-kt}\,. \qquad \text{(A)}$$

It remains to find the value of k. We know that the body cools to 60° in 20 min, i.e. that $T=60$ with $t=20$. This gives us

$$40=80e^{-20k},$$

whence we could in fact find k. Instead, we find e^{-k}. We have $e^{-20k}=\frac{40}{80}=\frac{1}{2}$ and $e^{-k}=\left(\frac{1}{2}\right)^{1/20}$.

We can now write equation (A) as

$$T-20=80\left(\frac{1}{2}\right)^{\frac{t}{20}}. \tag{B}$$

We have thus found the equation for the temperature T as a function of time t.

The problem asks us to find the value of t for $T=30$. We can do this with the aid of the above equation. On substituting $T=30$ in equation (B), we obtain

$$10=80\left(\frac{1}{2}\right)^{\frac{t}{20}},$$

whence $\left(\frac{1}{2}\right)^{t/20}=\frac{1}{8}=\left(\frac{1}{2}\right)^3$; thus $\frac{t}{20}=3$ and $t=60$.

The body thus cools to a temperature of 30° in the course of an hour.

4. Let 100 l. of salt solution contain 10 kg of salt. Water flows into the vessel containing the solution at the rate of 3 l. per min, and solution flows out at the same rate, the concentration being maintained uniform by stirring. How much salt will the solution in the vessel contain after an hour?

Solution. Let x be the amount of salt in the vessel after t min. Then the concentration will be

$$c=\frac{x}{100} \text{ kg per l.}$$

If the rate of increase in the amount of salt in the vessel at the instant t is given by the derivative $\frac{dx}{dt}$, the corresponding rate of decrease will be $-\frac{dx}{dt}$.

On the other hand, since the solution flows from the vessel at a rate of 3 l. per min, the rate of change in the amount of salt must be equal to $\frac{3x}{100}$. Hence

$$\frac{dx}{dt} = -\frac{3x}{100}.$$

We can rewrite this as

$$\frac{dx}{x} = -0{\cdot}03\,dt.$$

Integration gives us

$$\ln x = -0{\cdot}03t + C.$$

Since the amount of salt $x = 10$ with $t = 0$, this last equation gives us $\ln 10 = C$.

We have thus obtained the expression for the amount of salt after t min:

$$\ln x = -0{\cdot}03t + \ln 10.$$

We find the amount of salt remaining in the vessel after an hour by putting $t = 60$:

$$\ln x = -1{\cdot}8 + \ln 10.$$

We find from tables of natural logarithms that $x = 1{\cdot}654$ kg.

§ 94. **Differential equations with separable variables.** If a differential equation can be reduced by means of various transformations to the form

$$f(x)\,dx + \varphi(y)\,dy = 0, \tag{3}$$

we have an equation with *separated variables.* The process of reducing the equation to this form is called *separation of the variables.*

Our discussion of the integration of equation (3) will completely ignore the question of finding its singular solutions.

We obtain on integrating the left- and right-hand sides of equation (3):

$$\int f(x)\,dx+\int\varphi(y)\,dy=C.$$

This relationship is the *general* integral of equation (3).

If the relationship in question can be solved with respect to the function y, we can obtain the *general solution* of equation (3). The problem of integrating equation (3) is reckoned to be solved, however, as soon as we have obtained its general integral.

Equations with separable variables which are not given in as simple a form as (3) can be reduced to this form by using the rule given below for separating the variables.

First step. We reduce all the terms in the equation to a common denominator and cancel it out. If the equation contains derivatives, we multiply by the differential of the independent variable.

Second step. We collect all the terms containing the differential of the independent variable and similarly, all the terms containing the differential of the function. After this the equation takes the form

$$XY\,dx=X_1Y_1dy,$$

where X, X_1 are functions of the single variable x, and Y, Y_1 are functions of the single variable y.

Third step. On now dividing through by X_1Y, we arrive at an equation of the form (3).

Example. Integrate the equation

$$(1-x^2)\frac{dy}{dx}+xy=ax.$$

Solution.

First step:

$$(1-x^2)\,dy+xy\,dx=ax\,dx.$$

Second step:

$$x\,(y-a)\,dx+(1-x^2)\,dy=0.$$

Third step:

$$\frac{x\,dx}{1-x^2}+\frac{dy}{y-a}=0.$$

Fourth step:

$$\int\frac{x\,dx}{1-x^2}+\int\frac{dy}{y-a}=0.$$

whence
$$-\frac{1}{2}\ln(1-x^2)+\ln(y-a)=\ln c,$$

where C is written as $\ln c$.

Hence

$$\frac{y-a}{\sqrt{1-x^2}}=c, \quad y=a+c\sqrt{1-x^2}.$$

EXERCISES

Integrate the following equations:

1. $y\,dx-x\,dy=0.$ Ans. $y=cx.$
2. $(1+y)\,dx-(1-x)\,dy=0.$ Ans. $(1+y)(1-x)=c.$
3. $(1+x)\,y\,dx+(1-y)\,x\,dy=0.$ Ans. $\ln(xy)+x-y=c.$
4. $(x^2-yx^2)\dfrac{dy}{dx}+y^2+xy^2=0.$ Ans. $\dfrac{x+y}{xy}+\ln\dfrac{y}{x}=c.$
5. $x^2\,dy+(y-a)\,dx=0.$ Ans. $y-a=ce^{\frac{1}{x}}.$
6. $(1+y^2)\,dx-x^{\frac{1}{2}}\,dy=0.$ Ans. $2\sqrt{x}-\arctan y=c.$
7. $dy+y\tan x\,dx=0.$ Ans. $y=c\cos x.$
8. $\cos x\sin y\,dy-\cos y\sin x\,dx=0.$ Ans. $\cos y=c\cos x.$

9. Find the equation of the curve such that the tangent at any point of it of abscissa x has a slope of $3x-2$.

Ans. $y=\frac{3}{2}x^2-2x+C$.

10. Find the equation of the curve passing through the point $(0, 3)$ such that the slope of the tangent at any point of abscissa x is equal to x^2+5x.

Ans. $y=\frac{1}{3}x^3+\frac{5}{2}x^2+3$.

11. Find the equation of the curve through the point $(1, 1)$ such that the slope of the tangent at any point is proportional to the square of the ordinate of the point.

Ans. $k(x-1)y-y+1=0$, where k is a coefficient of proportionality.

Assuming that $s=0$ at $t=0$, find the relationships between s and t, given that the velocity v is equal to:

12. A constant v_0. Ans. $s=v_0 t$.

13. $m+kt$. Ans. $s=mt+\frac{1}{2}kt^2$.

14. $3+2t-3t^2$. A . $s=3t+t^2-t^3$.

15. A particle moving from rest has a velocity of $5t^2$ m/sec after t sec. Find: 1) how far it will have travelled from the point of departure after 3 sec; 2) how long it will take to travel 360 m from the point of departure.

Ans. 1) 45 m; 2) 6 sec.

16. Find the law of motion of a particle moving with a velocity proportional to the distance travelled, given that the particle has gone 100 m after 10 sec and 200 m after 15 sec.

Ans. $s=25\cdot(2)^{\frac{t}{5}}$

17. If the air temperature is 15°C and we know that a body cools from 90 to 40°C in 30 min, what will be the temperature of the body an hour after the first measurement?

Ans. $23^1/_3$°C.

Hint. The solution follows the lines of example 3, § 93.

18. The activity of a radioactive decay is proportional to its rate of decrease. Find the activity as a function of time, given that it decreases to a half in 4 days.

Ans. $R=R_0\, e^{-\frac{1}{4}(\ln 2)t}$

19. 100 l. of salt solution contains 10 kg. salt. Water enters the vessel containing the solution at 3 l. per min and the solution is run off at 2 l. per min, the concentration being kept uniform by stirring. How much salt will the vessel contain after 1 hour?

Ans. 3·9 kg.

CHAPTER 12

SERIES

§ 95. Numerical series. Fundamental concepts and theorems.

1. Definition. *A numerical series is an expression of the form*

$$a_1+a_2+a_3+\ldots+a_n+\ldots, \tag{1}$$

in which $a_1, a_2, a_3, \ldots, a_n, \ldots$ (the terms of the series) *are definite numbers, each term* a_n *being determined for a given n in accordance with some known law.*

Example 1. Let the series be

$$1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\ldots+\frac{1}{n}+\ldots$$

Here $a_n=\frac{1}{n}$. If we want to find say the twenty-third term, we set $n=23$ and find $a_{23}=\frac{1}{23}$; similarly, we have $a_{100}=\frac{1}{100}$, $a_{375}=\frac{1}{375}$ and so on.

Example 2. Consider the series

$$\frac{1}{1\,.\,2}+\frac{1}{2\cdot 3}+\frac{1}{3\,.\,4}+\ldots+\frac{1}{n\,(n+1)}+\ldots$$

Here $a_n=\frac{1}{n(n+1)}$. Hence $a_7=\frac{1}{7.8}=\frac{1}{56}$, $a_{30}=\frac{1}{30.31}=\frac{1}{930}$ and so on.

The expression giving the n-th term of series (1) for any value of $n=1, 2, 3, \ldots$ is called the *general term* of the series and is denoted by the symbol a_n. Thus the general term in our first example is $a_n=\frac{1}{n}$, whilst in our second example $a_n=\frac{1}{n\,(n+1)}$.

2. An example of an infinite series will already have been encountered in elementary mathematics, when dealing with the summation of the terms of the infinite decreasing geometrical progression.

$$a\,,\quad aq\,,\quad aq^2\,,\dots,\quad aq^{n-1}\,,\dots,$$

where $|q|<1$ and a is any real number.

We can form from the terms of this progression the infinite series

$$a+aq+aq^2+\dots+aq^{n-1}+\dots \tag{2}$$

The sum of the first n terms of this series is given, like the sum or the first n terms of the geometrical progression, by the familiaf expression

$$A_n=\frac{a-aq^n}{1-q}\,.$$

Since $|q|<1$, $\lim\limits_{n\to\infty} aq^n=0$, so that

$$\lim_{n\to\infty} A_n=\frac{a}{1-q}\,.$$

This limit is called the sum of the geometrical progression and is also the sum of series (2).

3. As we have seen, the sum of the geometrical progression (i.e. the sum of series (2)) is found by forming the sum A_n of its first n terms then proceeding to the limit of the sum as $n\to\infty$. An extension of this procedure leads to the idea of the sum of a general numerical series.

We put

$$A_n=a_1+a_2+a_3+\dots+a_n\,.$$

As n passes through the values 1, 2, 3, ..., A_n takes the sequence of values $A_1=a_1$, $A_2=a_1+a_2$, $A_3=a_1+a_2+a_3$, ...

Thus A_n is a variable that depends on n (we can look on it as a function of n). The numbers A_1, A_2, A_3, ... are known as *partial* sums of series (1), A_n being known as the n-th partial sum of the series. If the limit exists, i.e.

$$\lim_{n\to\infty} A_n = A\,, \tag{3}$$

series (1) is said to *converge* to the *sum* A and we write

$$A = a_1 + a_2 + \ldots + a_n + \ldots$$

or

$$A = \sum_{n=1}^{\infty} a_n\,.$$

On the other hand, if the limit (3) does not exist, we say that series (1) is *divergent.*

Example 1. Consider the series

$$\frac{1}{1\,.\,2} + \frac{1}{2\,.\,3} + \frac{1}{3\,.\,4} + \ldots + \frac{1}{n\,(n+1)} + \ldots$$

On observing that $\frac{1}{n\,(n+1)} = \frac{1}{n} - \frac{1}{n+1}$, we can write each term of the series as a difference, viz.

$$a_1 = \frac{1}{1\,.\,2} = \frac{1}{1} - \frac{1}{\ }\,,\quad a_2 = \frac{1}{2\,.\,3} = \frac{1}{2} - \frac{1}{3}\,,\quad a_3 = \frac{1}{3} - \frac{1}{4}\,, \ldots,$$

whilst its partial sum A_n may be written as

$$A_n = a_1 + a_2 + \ldots + a_n$$
$$= \left(\frac{1}{1} - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \left(\frac{1}{3} - \frac{1}{4}\right) + \ldots + \left(\frac{1}{n} - \frac{1}{n+1}\right).$$

We cancel and obtain

$$A_n = 1 - \frac{1}{n+1}\,;$$

whence

$$\lim_{n\to\infty} A_n = \lim_{n\to\infty}\left(1 - \frac{1}{n+1}\right) = 1.$$

The series is thus convergent and its sum is 1.

Example 2. Let the given series be

$$\ln\frac{2}{1}+\ln\frac{3}{2}+\ln\frac{4}{3}+\ldots+\ln\frac{n+1}{n}+\ldots$$

Here $a_n=\ln\frac{n+1}{n}=\ln(n+1)-\ln n$. Hence

$$A_n=(\ln 2-\ln 1)+(\ln 3-\ln 2)+(\ln 4-\ln 3)+\ldots$$
$$+(\ln(n+1)-\ln n)=\ln(n+1).$$

Thus

$$\lim_{n\to\infty} A_n=\lim_{n\to\infty}\ln(n+1)=+\infty,$$

whence we conclude that the series is divergent.

Example 3. We again consider series (2)

$$a+aq+aq^2+\ldots+aq^{n-1}+\ldots,$$

formed from the terms of the geometrical progression

$$a,\quad aq,\quad aq^2,\ldots,\quad aq^{n-1},\ldots$$

In mathematical analysis, series (2) is also described as a geometrical progression, and we shall adopt this terminology in future. If $|q|<1$, progression (2) converges, as we have seen, and its sum is $\frac{a}{1-q}$. We now take the case of $|q|>1$. We have $\lim aq^n=\infty$, so that $\lim\limits_{n\to\infty} A_n=\lim\limits_{n\to\infty}\frac{a-aq^n}{1-q}=\infty$.
Hence (2) diverges for $|q|>1$.

If $q=1$, the partial sum A_n of series (2) is equal to na and $\lim\limits_{n\to\infty} A_n=\lim\limits_{n\to\infty}(na)=\infty$, i.e. the series diverges.

Finally, if $q=-1$, series (2) becomes

$$a-a+a-a+\ldots \tag{4}$$

Thus $A_1=a$, $A_2=0$, $A_3=a$, $A_4=0$, ..., and in general, $A_n=a$ if n is odd and $A_n=0$ with n even. We see that, while the partial sum is bounded it does not tend to a limit, so that series (4) is divergent.

Hence *the geometrical progression* (2) *is convergent if and only if* $|q|<1$, *in which case its sum* $A=\frac{a}{1-q}$.

Example 4. The so-called harmonic series*

$$1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{5}+\dots, \tag{5}$$

plays an important part in the theory of series.

We show that the harmonic series is divergent, i.e. that $\lim_{n\to\infty} A_n = +\infty$. This is done by considering the following partial sums of series (5):

$$A_2=1+\frac{1}{2}; \quad A_4=A_2+\frac{1}{3}+\frac{1}{4};$$

$$A_8=A_4+\frac{1}{5}+\frac{1}{6}+\frac{1}{7}+\frac{1}{8},$$

$$A_{16}=A_8+\frac{1}{9}+\frac{1}{10}+\frac{1}{11}+\frac{1}{12}+\frac{1}{13}+\frac{1}{14}+\frac{1}{15}+\frac{1}{16}.$$

It is easily seen that

$$A_2 > \frac{1}{2}+\frac{1}{2}=2\cdot\frac{1}{2},$$

$$A_4 > \frac{1}{2}+\frac{1}{2}+\frac{1}{4}+\frac{1}{4}=\frac{1}{2}+\frac{1}{2}+\frac{1}{2}=3\cdot\frac{1}{2},$$

$$A_8 > \frac{1}{2}+\frac{1}{2}+\frac{1}{2}+\frac{1}{8}+\frac{1}{8}+\frac{1}{8}+\frac{1}{8}$$

* The number c is termed the harmonic mean of the numbers a and b if

$$\frac{1}{c}=\frac{1}{2}\left(\frac{1}{a}+\frac{1}{b}\right).$$

It may easily be seen that, starting with the second, each term of series (5) is the harmonic mean of the two neighbouring terms. Hence the name "harmonic series".

$$=\frac{1}{2}+\frac{1}{2}+\frac{1}{2}+\frac{1}{2}=4\cdot\frac{1}{2},$$

$$A_{16} > \frac{1}{2}+\frac{1}{2}+\frac{1}{2}+\frac{1}{2}+\frac{1}{16}+\frac{1}{16}+\frac{1}{16}+\frac{1}{16}+\frac{1}{16}+\frac{1}{16}+$$
$$+\frac{1}{16}+\frac{1}{16}=\frac{1}{2}+\frac{1}{2}+\frac{1}{2}+\frac{1}{2}+\frac{1}{2}=5\cdot\frac{1}{2}.$$

We find similarly that

$$A_{32} > 6\cdot\frac{1}{2}, \quad A_{64} > 7\cdot\frac{1}{2}.$$

Hence it is clear that the partial sum with subscript $m=2^k$ $(k=1, 2, 3, \ldots)$ increases indefinitely as m runs through the values 2, 4, 8, 16, 32, 64, ... And since the condition $A_1 < A_2 < A_3 < A_4 < A_5 < < A_6 < A_7 < \ldots$ is fulfilled for the series, we conclude that

$$\lim_{n\to\infty} A_n = +\infty \quad (n=1, 2, 3, \ldots).$$

4. When the expression

$$a_1+a_2+\ldots+a_n+\ldots \tag{1}$$

represents a convergent series, the numerical meaning of the expression is conveyed by the fact of its having the sum A, i.e.

$$a_1+a_2+\ldots+a_n+\ldots=A.$$

Expression (1) has no numerical meaning when the series is divergent.

5. THEOREM. *The general term a_n of a convergent series* (1) *tends to zero as $n\to\infty$.*

Proof. Since

$$A_n=a_1+a_2+\ldots+a_{n-1}+a_n$$

and

$$A_{n-1}=a_1+a_2+\ldots+a_{n-1},$$

we have

$$a_n=A_n-A_{n-1}.$$

The theorem tells us that series (1) is convergent, whence it fol-

lows that

$$\lim_{n\to\infty}(A_n - A_{n-1}) = \lim_{n\to\infty} A_n - \lim_{n\to\infty} A_{n-1} = A - A = 0. \qquad (5/10)$$

Thus we have $\lim\limits_{n\to\infty} a_n = 0$.

It must be particularly emphasized that the above theorem expresses only a *necessary* condition for the convergence of the series. This means that, if series (1) is known to be convergent, we can rightly conclude that $\lim\limits_{n\to\infty} a_n = 0$; whereas if we are only told that $\lim\limits_{n\to\infty} a_n = 0$ and have no idea whether the series is convergent or not, we cannot conclude that the series is in fact convergent, since it may quite well be divergent. For instance, the general term $\frac{1}{n}$ of the harmonic series (5) tends to 0 as $n\to\infty$, although the harmonic series is divergent. On the other hand, if a_n does not tend to zero as $n\to\infty$, the series must certainly be divergent (if it were convergent, its general term a_n would tend to 0).

§ **96. Series with positive terms.** 1. Series in which all the terms are positive numbers possess certain special features which have contributed to the considerable practical importance of such series. Apart from the fact that they are easy to investigate in general, it has been possible to devise special tests whereby the question of the convergence or divergence of the series can be decided very simply. We shall consider two of the simplest of these tests.

2. *Comparison test for series.*

Given the two series with positive terms:

and $$a_1 + a_2 + \ldots + a_n + \ldots \qquad (A)$$

let $$b_1 + b_2 + \ldots + b_n + \ldots, \qquad (B)$$

$$a_1 \leqslant b_1,\quad a_2 \leqslant b_2,\ \ldots,\ a_n \leqslant b_n,\ \ldots$$

If series (B) *is convergent, then series* (A) *is also convergent* (*and consequently, if series* (A) *is divergent, series* (B) *is also divergent*).

Proof. Let A_n, B_n denote respectively the sums of the first n terms of series (A) and (B). If series (B) is convergent, the limit exists, i.e.

$$\lim_{n\to\infty} B_n = B,$$

where B is a positive number.

Since all the terms of series (B) are positive, we have

$$B_1 < B_2 < B_3 < \dots < B_n < \dots$$

and clearly, $B_n < B$. It follows from the relationships

$$a_1 \leqslant b_1\,, \quad a_2 \leqslant b_2\,, \dots, \quad a_n \leqslant b_n\,, \dots$$

that $A_n \leqq B_n$, so that $A_n < B$. As in the case of series (B), we have

$$A_1 < A_2 < \dots < A_n < \dots$$

FIG. 113

We shall represent the sums $A_1, A_2, \dots, A_n, \dots$ as points on a number axis (fig. 113). Since all the numbers $A_1, A_2, \dots, A_n, \dots$ remain less than B and at the same time $A_1 < A_2 < A_3 < \dots$, the corresponding points must accumulate about some point A, either lying to the left of or coinciding with B (fig. 113). If we take an interval of length ε, no matter how small, to the left of the point A, then after a sufficiently large n every point A_n will lie inside this interval (since each successive point lies to the right of the one before) and will remain in the interval. Hence, as from a certain n, the inequality

will hold:

$$|A_n - A| < \varepsilon.$$

Thus it follows that $\lim\limits_{n\to\infty} A_n = A$. But if $\lim\limits_{n\to\infty} A_n$ exists, series (A) is convergent.

It follows at once from what has been proved that, if series (A) is divergent, series (B) is also divergent (because, if this were not so, series (A) would be convergent).

Example 1. We consider the series

$$1+\frac{1}{2^2}+\frac{1}{3^3}+\frac{1}{4^4}+\ldots+\frac{1}{n^n}+\ldots \tag{6}$$

On neglecting the first term, we have the series

$$\frac{1}{2^2}+\frac{1}{3^3}+\frac{1}{4^4}+\ldots+\frac{1}{n^n}+\frac{1}{(n+1)^{n+1}}+\ldots \tag{6*}$$

We compare this series with the convergent geometrical progression

$$\frac{1}{2^2}+\frac{1}{2^3}+\frac{1}{2^4}+\ldots+\frac{1}{2^n}+\frac{1}{2^{n+1}}+\ldots$$

We have

$$\frac{1}{2^2}=\frac{1}{2^2}; \quad \frac{1}{3^3}<\frac{1}{2^3}; \quad \frac{1}{4^4}<\frac{1}{2^4}; \ldots$$

Consequently series (6*) is convergent, and so, clearly, is series (6), its sum being equal to 1 plus the sum of series (6*).

Example 2. Let the given series be

$$1+\frac{1}{\sqrt{2}}+\frac{1}{\sqrt{3}}+\frac{1}{\sqrt{4}}+\ldots$$

Since

$$\frac{1}{\sqrt{n}} \geq \frac{1}{n} \quad (n=1, 2, 3, \ldots),$$

and $\frac{1}{n}$ is the general term of the divergent (harmonic) series

$$1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\ldots,$$

the given series must be divergent.

Example 3. We take the series

$$\frac{1}{1^2}+\frac{1}{2^2}+\frac{1}{3^2}+\frac{1}{4^2}+\ldots+\frac{1}{n^2}+\ldots \tag{7}$$

If we can show that the series

$$\frac{1}{2^2}+\frac{1}{3^3}+\frac{1}{4^2}+\ldots+\frac{1}{n^2}+\frac{1}{(n+1)^2}+\ldots, \tag{8}$$

is convergent, it will follow that the original series is also convergent.

We compare series (8) with

$$\frac{1}{1.2}+\frac{1}{2.3}+\frac{1}{3.4}+\ldots+\frac{1}{n(n+1)}+\ldots,$$

which we know to be convergent from § 95. Since

$$\frac{1}{(n+1)^2}<\frac{1}{n(n+1)},$$

we conclude from the comparison test that series (8) is convergent; hence series (7) is also convergent.

Remark. It may easily be shown that the comparison test remains valid for series with zero as well as positive terms, series of this type being described as non-negative.

3. d'Alembert's convergence test for series with positive terms.

THEOREM. *Let all the terms of the series*

$$a_1+a_2+\ldots+a_n+\ldots \tag{1}$$

be positive, and let the ratio of the $(n+1)$-th *term to the* n-th *term have a limit equal to* l *as* n *increases indefinitely,* i.e.

$$\lim_{n\to\infty}\frac{a_{n+1}}{a_n}=l.$$

Then:

1) *If the limit* $l<1$, *series* (1) *is convergent.*

2) *If* $l>1$, *series* (1) *is divergent.*

3) If $l=1$, the test does not give a definite answer, some series being convergent in this case and others divergent.

Proof. 1) $l<1$. We take any number q greater than l and less than 1 $(l<q<1)$. Since the ratio $\frac{a_{n+1}}{a_n}\to l$ as $n\to\infty$, there

must be a certain $n=m$ after which the ratio remains less than q as n increases. Hence we have

$$\frac{a_{m+1}}{a_m}<q\,, \quad \frac{a_{m+2}}{a_{m+1}}<q\,, \quad \frac{a_{m+3}}{a_{m+2}}<q\,, \ldots$$

It follows that $a_{m+1}<a_mq$, $a_{m+2}<a_{m+1}q$, and since $a_{m+1}<$ $<a_mq$, it follows all the more that $a_{m+2}<a_mq^2$. Similarly, $a_{m+3}<a_{m+2}q<a_mq^3$, $a_{m+4}<a_{m+3}q<a_mq^4$, . . .

Thus

$$\left.\begin{array}{l} a_{m+1}<a_mq\,, \\ a_{m+2}<a_mq^2\,, \\ a_{m+3}<a_mq^3\,, \\ a_{m+4}<a_mq^4\,, \\ \ldots \end{array}\right\} \tag{9}$$

We now compare the series

$$a_{m+1}+a_{m+2}+a_{m+3}+\ldots \tag{10}$$

and

$$a_mq+a_mq^2+a_mq^3+\ldots \tag{11}$$

Series (11) is a progression in which q is a positive number less than 1, and so series (11) is convergent. Using relationships (9), the comparison test shows that series (10) must also be convergent. Let its sum be A^*, i.e.

$$a_{m+1}+a_{m+2}+a_{m+3}+\ldots=A^*\,. \tag{12}$$

The orginal series (1), i.e. the series

$$a_1+a_2+a_3+\ldots+a_m+a_{m+1}+a_{m+2}+\ldots$$

differs from (12) only in the fact that it contains the m extra terms $a_1, a_2, \ldots, a_m$. The sum of these terms is some definite number A_m. Consequently

$$a_1+a_2+a_3+\ldots+a_m+a_{m+1}+a_{m+2}+\ldots=A_m+A^*=A\,,$$

i.e. series (1) is convergent.

2) $l>1$. Since in this case the ratio $\frac{a_{n+1}}{a_n}$ tends to a number $l>1$ as $n\to\infty$, it will remain greater than 1 after some value of $n=m$ say, i.e.

$$\frac{a_{m+1}}{a_m}>1, \quad \frac{a_{m+2}}{a_{m+1}}>1, \quad \frac{a_{m+3}}{a_{m+2}}>1, \ldots$$

Hence we have

$$a_{m+1}>a_m,$$

$$a_{m+2}>a_{m+1}>a_m,$$

$$a_{m+3}>a_{m+2}>a_m,$$

The term a_m of series (1) is a positive number, and in view of the fact that all the terms of (1), starting from a_{m+1}, remain greater than a_m, the general term of (1) cannot tend to zero as $n\to\infty$. Thus series (1) must be divergent: for if it were convergent, its general term would tend to zero as $n\to\infty$ in accordance with the necessary condition for convergence (§ 95, 5), and this is not the case.

3) $l=1$. We shall use examples to show the truth of the statement contained in the theorem.

We know that the series

$$\frac{1}{1\cdot 2}+\frac{1}{2\cdot 3}+\frac{1}{3\cdot 4}+\ldots$$

is convergent (see § 95, 3, example 1). Here

$$a_n=\frac{1}{n(n+1)}, \qquad a_{n+1}=\frac{1}{(n+1)(n+2)}$$

and $$\lim_{n\to\infty}\frac{a_{n+1}}{a_n}=\lim_{n\to\infty}\frac{n(n+1)}{(n+1)(n+2)}=\lim_{n\to\infty}\frac{n}{n+2}$$

$$=\lim_{n\to\infty}\frac{1}{\frac{n+2}{n}}=\lim_{n\to\infty}\frac{1}{1+\frac{2}{n}}=1.$$

The harmonic series

$$1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\dots$$

is divergent. At the same time

$$\lim_{n\to\infty}\frac{a_{n+1}}{a_n}=\lim_{n\to\infty}\frac{n}{n+1}=\lim_{n\to\infty}\frac{1}{\frac{n+1}{n}}=\lim_{n\to\infty}\frac{1}{1+\frac{1}{n}}=1.$$

We conclude by considering two examples of the application of d'Alembert's test.

1. Let the given series be

$$1+\frac{2}{3}+\frac{3}{3^2}+\frac{4}{3^3}+\frac{5}{3^4}+\dots$$

Here
$$a_n=\frac{n}{3^{n-1}},\qquad a_{n+1}=\frac{n+1}{3^n},$$

$$\frac{a_{n+1}}{a_n}=\frac{(n+1)\,3^{n-1}}{3^n\cdot n}=\frac{1}{3}\left(1+\frac{1}{n}\right)$$

and
$$\lim_{n\to\infty}\frac{a_{n+1}}{a_n}=\lim_{n\to\infty}\left[\frac{1}{3}\left(1+\frac{1}{n}\right)\right]=\frac{1}{3}<1.$$

The series is thus convergent.

2. We take the series

$$1+\frac{2^2}{2!}+\frac{3^3}{3!}+\frac{4^4}{4!}+\dots^{*}$$

We have

$$a_n=\frac{n^n}{n!},\qquad a_{n+1}=\frac{(n+1)^{n+1}}{(n+1)!}$$

$$\frac{a_{n+1}}{a_n}=\frac{(n+1)^{n+1}\cdot n!}{(n+1)!\,n^n}=\frac{(n+1)^n}{n^n}=\left(\frac{n+1}{n}\right)^n=\left(1+\frac{1}{n}\right)^n.$$

* The symbol $n!$ (factorial n) stands for the product 1. 2. 3. . . . n. Thus $1!=1$. $2!=1.2=2$, $3!=1.2.3=6$, $4!=1.2.3.4=24$, $5!=1.2.3.4.5=120$ and so on.

Consequently $$\lim_{n\to\infty}\frac{a_{n+1}}{a_n}=\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n=e\ .$$

Since $e=2{\cdot}71828\ldots>1$, the series is divergent.

§ 97. **Alternating series.** An alternating series is one in which the terms are alternately positive and negative. For instance, the series

$$1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\ldots \tag{13}$$

is alternating.

The following theorem applies for alternating series.

Theorem (Leibniz). *If the terms of an alternating series decrease in absolute value and tend to zero as $n\to\infty$, the series is convergent.*

We omit the proof, which is outside the scope of an elementary course.

The conditions of the theorem are satisfied by series (13), for

$$1>\frac{1}{2}>\frac{1}{3}>\frac{1}{4}>\ldots>\frac{1}{n}>\ldots\ ;$$

and the terms are thus decreasing in absolute value. Moreover the general term a_n is of the form $a_n=(-1)^{n-1}\cdot\frac{1}{n}$ and $\lim\limits_{n\to\infty}(-1)^{n-1}\cdot\frac{1}{n}=0$, so that the second condition is satisfied. Hence series (13) is convergent.

§ 98. **Absolute convergence.** We shall now consider series with arbitrary terms, i.e. those containing an infinite set of positive and an infinite set of negative terms (zero terms may be included).

THEOREM. *Let the series*

$$a_1+a_2+\ldots+a_n+\ldots \tag{1}$$

have arbitrary terms. If the series composed of the absolute values

of the terms of series (1) *is convergent,* i.e. *the series*

$$|a_1|+|a_2|+\ldots+|a_n|+\ldots, \tag{14}$$

is convergent, then (1) *is also convergent.*

Proof. We form the auxiliary series

$$(a_1+|a_1|)+(a_2+|a_2|)+(a_3+|a_3|)+\ldots \tag{15}$$

If $a_n \geqq 0$, we have $|a_n|=a_n$; whilst if $a_n<0$, $|a_n|=-a_n$. In the first case $a_n+|a_n|=2a_n$, and in the second $a_n+|a_n|=0$. Thus each term of series (15) is equal to twice the corresponding term of series (14) or else it is zero.

Series (14) is convergent by hypothesis. Then the series

$$2|a_1|+2|a_2|+2|a_3|+\ldots \tag{14*}$$

is also convergent.

For, if

$$|a_1|+|a_2|+\ldots+|a_n|=A_n,$$

then
$$2|a_1|+2|a_2|+\ldots+2|a_n|=2A_n.$$

Since series (14) is convergent, there exists

$$\lim_{n\to\infty} A_n=A.$$

Thus
$$\lim_{n\to\infty}(2A_n)=2\lim_{n\to\infty} A_n=2A,$$

and series (14*) is therefore convergent.

Since $2|a_n| \geqq a_n+|a_n|$, it follows that series (15) is convergent (by the comparison theorem for positive series).

We put

$$A_n=a_1+a_2+a_3+\ldots+a_n,$$

$$A_n^*=|a_1|+|a_2|+|a_3|+\ldots+|a_n|,$$

$$A_n^{**}=(a_1+|a_1|)+(a_2+|a_2|)+\ldots+(a_n+|a_n|).$$

Hence

$$A_n^{**}-A_n^*=a_1+a_2+\ldots+a_n,$$

i.e.

$$A_n^{**}-A_n^*=A_n.$$

Series (14) is convergent by hypothesis, whilst we have established the convergence of series (15). Consequently the limits $\lim_{n\to\infty} A_n^{**}$ and $\lim_{n\to\infty} A_n^{*}$ exist; the limit $\lim_{n\to\infty} (A_n^{**} - A_n^{*})$ therefore exists, so that $\lim_{n\to\infty} A_n$ must exist; hence series (1) is convergent.

Series (1), for which the conditions of the last theorem are satisfied, is said to be *absolutely convergent*. Thus every absolutely convergent series is also convergent.

We return to series (13), i.e.

$$1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\ldots$$

By Leibniz's theorem, this series is convergent. We form the series from the absolute values of its terms, i.e.

$$1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\ldots$$

This last is the harmonic series and diverges. Thus series (13) is convergent, but not absolutely convergent.

This means that there can be both absolutely convergent and non-absolutely convergent series.

Absolutely convergent series have certain important properties not possessed by non-absolutely convergent series. Due to these properties, absolutely convergent series are of great importance in pure and applied mathematics.

Example. We show that the series

$$\frac{\sin\alpha}{2}+\frac{\sin 2\alpha}{4}+\frac{\sin 3\alpha}{8}+\frac{\sin 4\alpha}{16}+\ldots \tag{16}$$

is convergent (and converges absolutely).

We form the series

$$\frac{|\sin\alpha|}{2}+\frac{|\sin 2\alpha|}{4}+\frac{|\sin 3\alpha|}{8}+\frac{|\sin 4\alpha|}{16}+\ldots \tag{17}$$

For series (17), $a_n=\dfrac{|\sin n\alpha|}{2^n}$; since $|\sin n\alpha|\leqq 1$, the terms of (17) do not exceed the corresponding terms of the convergent

progression

$$\frac{1}{2}+\frac{1}{2^2}+\frac{1}{2^3}+\dots$$

By the comparison theorem for non-negative series, series (17) converges. Hence series (16) is convergent, and converges absolutely.

§ **99. Functional series.** We consider the expression

$$x+x^2+x^3+\dots+x^n+\dots, \tag{18}$$

where x is an independent variable. If we assign the value $\frac{1}{2}$ to x, expression (18) becomes the numerical series

$$\frac{1}{2}+\frac{1}{2^2}+\frac{1}{2^3}+\dots \tag{18*}$$

This is the geometrical progression in which $q=\frac{1}{2}$, so that (18*) is convergent. On putting $x=2$, we obtain the numerical series

$$2+2^2+2^3+\dots,$$

which is clearly divergent.

The terms $x, x^2, x^3, \dots$ which make up series (18) are functions of the argument x. Hence (18) is called a *functional* series. Given a definite numerical value of x, (18) becomes an ordinary numerical series. These latter may be convergent for certain values of x and divergent for others, as we have seen. The set of *all* values of x for which series (18) is convergent is called the *domain of convergence* of the series, whilst the *domain of divergence* of (18) is the set of all values of x for which it diverges. The domain of convergence of (18) is easily seen to be the interval $(-1, 1)$; for (18) is a progression in x, and a progression is only convergent for $|x|<1$. The domain of divergence of the series is the set of values of x satisfying $|x|\geqq 1$, or $-\infty<x\leqq -1$ and $1\leqq x<+\infty$.

If x lies in the (open) interval $(-1, 1)$, series (18) has the sum $\frac{x}{1-x}$ which depends on x, i.e. is a function of x.

In general, a functional series is an expression of the form

$$u_1(x)+u_2(x)+\ldots+u_n(x)+\ldots, \tag{19}$$

where $u_1(x)$, $u_2(x)$, ... are functions of the same variable x defined in a certain domain. The set of values of x for which series (19) converges (becoming a numerical series for each x) is termed the *domain of convergence* of the series. The sum of series (19) is obviously a function of x.

The most important types of functional series include power and trigonometric (Fourier) series, which we shall discuss in the rest of this chapter.

§ 100. **Power series.** A power series is defined as a functional series of the form

$$a_0+a_1x+a_2x^2+a_3x^3+\ldots+a_nx^n+\ldots, \tag{20}$$

where x is the independent variable and $a_0, a_1, a_2, \ldots, a_n, \ldots$ are real numbers called the coefficients of the series.

The more general type of power series

$$a_0+a_1(x-\alpha)+a_2(x-\alpha)^2+\ldots+a_n(x-\alpha)^n+\ldots$$

is sometimes considered, where α is a real number. This type reduces to the form (20) with the aid of the substitution

$$x-\alpha=x'$$

so that it is permissible to confine ourselves to series (20).

The power series is the simplest type of functional series from the point of view of structure (its terms being the elementary power functions); hence series of this type are widely used both in analysis and applied mathematics.

All the terms of series (20) except the first vanish at $x=0$; hence the sum of any series of the form (20) exists at $x=0$ and is equal to a_0; in other words, (20) is convergent at $x=0$.

We next discuss the form of the domain of convergence of a power series converging for values of x other than $x=0$.

We start by finding the values of x for which series (20) is absolutely convergent.

We form a series from the absolute values of the terms of series (20), viz.,

$$|a_0|+|a_1 x|+|a_2 x^2|+\dots+|a_n x^n|+\dots \tag{21}$$

We write $|a_n x^n|=v_n$ and consider the ratio of the $(n+1)$-th term of series (21) to the n-th term, i.e.

$$\frac{v_{n+1}}{v_n}=\left|\frac{a_{n+1}x^{n+1}}{a_n x^n}\right|=\left|\frac{a_{n+1}}{a_n}\right||x|.$$

We assume $x\neq 0$ and let the limit of this ratio exist as $n\to\infty$. Since x is independent of n, we have

$$\lim_{n\to\infty}\frac{v_{n+1}}{v_n}=|x|\lim\left|\frac{a_{n+1}}{a_n}\right|.$$

We put

$$\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right|=l\neq 0.$$

Then

$$|x|\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right|=|x|l.$$

By d'Alembert's test, series (21) is convergent if the limit obtained is less than 1, i.e. if

$$|x|l<1,$$

or

$$|x|<\frac{1}{l},$$

or

$$-\frac{1}{l}<x<\frac{1}{l}.$$

Setting $\frac{1}{l}=R$, the last double inequality becomes

$$-R<x<R.$$

Series (21) is thus convergent in the interval $(-R, R)$ with centre at the point O (fig. 114). It follows that series (20) is convergent in this interval, and moreover converges absolutely (see § 98).

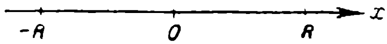


FIG. 114

By d'Alembert's test, series (21) will diverge for values of $|x|$ for which

$$|x| l > 1,$$

i.e. for x satisfying the inequality

$$|x| > \frac{1}{l},$$

or $|x| > R$; in other words, series (21) diverges outside the interval $(-R, R)$.

We proved when deducing d'Alembert's test (§ 96) that, if the limit of the ratio of any term to the previous term is greater than 1 as $n \to \infty$, the general term does not tend to zero as $n \to \infty$. Thus the general term $|a_n x^n|$ of series (21) does not tend to zero as $n \to \infty$ with $|x| > R$, so that the same is true for the general term $a_n x^n$ of series (20).

This means that the *necessary* condition for convergence is not fulfilled for series (20) with $|x| > R$ (§ 95), so that in this case series (20) is divergent.

We have now established that, if

$$\lim_{n \to \infty} \left| \frac{a_{n+1}}{a_n} \right| = l \neq 0,$$

power series (20) is absolutely convergent in the interval $(-R, R)$ and diverges outside it.

The number $R = \frac{1}{l}$ is called the *radius of convergence*, whilst $(-R, R)$ is the *interval of convergence* of the power series. We

have now settled the question of the convergence of series (20) except at the points $x=R$ and $x=-R$. As we shall see in later examples, a power series may be convergent or may be divergent for both these values of x, or it may converge for one value and diverge for the other.

We have found the form of the domain of convergence of a power series on the assumption that there exists $\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = l\neq 0$. The same result may be obtained by more subtle methods, to include the case when the limit does not always exist.

If, for any value of x,

$$\lim_{n\to\infty}\left(\left|\frac{a_{n+1}}{a_n}\right||x|\right)=0,$$

by d'Alembert's test, series (20) will converge (and converge absolutely) for any value of x, i.e. in the interval $(-\infty, +\infty)$ (over the entire numeral axis). In this case series (20) is said to have an infinite radius of convergence ($R=+\infty$).

Finally, with $x\neq 0$, let

$$\lim_{n\to\infty}\left(\left|\frac{a_{n+1}}{a_n}\right||x|\right)=+\infty\ ;$$

then series (20) can only converge for $x=0$ (the radius of convergence $R=0$). Series of this type are of no practical interest.

Example 1. Find the interval of convergence of the series

$$\frac{x}{1^2}+\frac{x^2}{2^2}+\frac{x^3}{3^2}+\ldots+\frac{x^n}{n^2}+\ldots \tag{22}$$

and determine whether or not it converges at the ends of the interval.

Solution. Here $a_0=0$, $a_n x^n=\frac{x^n}{n^2}$, $a_{n+1}x^{n+1}=\frac{x^{n+1}}{(n+1)^2}$.

We have

$$\lim_{n\to\infty}\left|\frac{a_{n+1}x^{n+1}}{a_n x^n}\right| = \lim_{n\to\infty}\frac{|x^{n+1}|\cdot n^2}{(n+1)^2\,|x^n|}$$

$$= \lim_{n\to\infty}\left[\left(\frac{n}{n+1}\right)^2 |x|\right] = \lim_{n\to\infty}\left[\frac{1}{\left(1+\frac{1}{n}\right)^2}|x|\right] = |x|.$$

We see from this that series (22) converges for $|x|<1$, i.e. in the interval $(-1, 1)$. With $x=-1$ series (22) becomes the numerical series

$$-\frac{1}{1^2}+\frac{1}{2^2}-\frac{1}{3^2}+\frac{1}{4^2}-\dots,$$

which is convergent by the theorem regarding alternating series (§ 97). With $x=1$ we get the series

$$\frac{1}{1^2}+\frac{1}{2^2}+\frac{1}{3^2}+\dots,$$

which we found to be convergent in § 96 (2, example 3). The series is thus convergent at both ends of the interval of convergence, i.e. its domain of convergence is the closed interval $[-1, +1]$.

Example 2. Find the interval of convergence of the series

$$\frac{x}{3}+\frac{x^3}{3^2}+\frac{x^5}{3^3}+\dots+\frac{x^{2n-1}}{3^n}+\dots \tag{23}$$

Here $a_0=a_2=a_4=\dots=0$. This fact does not prevent us from finding the interval of convergence in the same way as before, i.e. with the aid of d'Alembert's test.

We find, in fact, the limit of the absolute value of the ratio $\frac{u_{n+1}(x)}{u_n(x)}$ as $n\to\infty$ (i.e. the ratio of the $(n+1)$-th to the n-th term). We have

$$\lim_{n\to\infty}\left|\frac{u_{n+1}(x)}{u_n(x)}\right| = \lim_{n\to\infty}\frac{|x^{2n+1}|\,3^{n-1}}{3^n\,|x^{2n-1}|} = \frac{x^2}{3}.$$

The series will be convergent if $\frac{x^2}{3}<1$, i.e. if $|x|<\sqrt{3}$, so that

x must lie inside the interval $(-\sqrt{3}, +\sqrt{3})$. If $\frac{x^2}{3} > 1$, i.e. if $|x| > \sqrt{3}$, x lies outside this interval and series (23) is not absolutely convergent. But we proved above that, with $\lim_{n\to\infty}\left|\frac{u_{n+1}(x)}{u_n(x)}\right| > 1$, the general term of the series formed from the absolute values of the terms of (23) does not tend to zero as $n\to\infty$; in which case the general term $\frac{x^{2n-1}}{3^n}$ of the original series likewise does not tend to zero; thus series (23) is divergent for $|x| > \sqrt{3}$.

Further, with $x = -\sqrt{3}$ series (23) becomes the numerical series

$$-\frac{\sqrt{3}}{3}-\frac{3\sqrt{3}}{3^2}-\frac{3^2\sqrt{3}}{3^3}-\frac{3^3\sqrt{3}}{3^4}-\dots,$$

which is equivalent to

$$-\frac{\sqrt{3}}{3}-\frac{\sqrt{3}}{3}-\frac{\sqrt{3}}{3}-\frac{\sqrt{3}}{3}-\dots,$$

this being clearly divergent. With $x = +\sqrt{3}$ series (23) becomes the divergent numerical series

$$\frac{\sqrt{3}}{3}+\frac{3\sqrt{3}}{3^2}+\frac{3^2\sqrt{3}}{3^3}+\dots$$

Here we have an example of a series divergent at the ends of the interval of convergence.

Example 3. Find the interval of convergence of the series

$$\frac{x}{1}+\frac{x^2}{\sqrt{2}}+\frac{x^3}{\sqrt{3}}+\dots+\frac{x^n}{\sqrt{n}}+\dots \tag{24}$$

and determine whether or not it converges at the ends of the interval.

Solution. We have

$$\lim_{n\to\infty}\left|\frac{u_{n+1}(x)}{u_n(x)}\right| = \lim_{n\to\infty}\frac{|x^{n+1}|\sqrt{n}}{\sqrt{n+1}\,|x^n|} = \lim_{n\to\infty}\frac{|x|}{\sqrt{\frac{n+1}{n}}}$$

$$= \lim_{n\to\infty} \frac{|x|}{\sqrt{1+\frac{1}{n}}} = |x|.$$

The series is therefore convergent for $|x|<1$, i.e. in the interval $(-1, +1)$. With $x=-1$, the series becomes the numerical series

$$-1+\frac{1}{\sqrt{2}}-\frac{1}{\sqrt{3}}+\frac{1}{\sqrt{4}}-\ldots,$$

which is convergent by the theorem for alternating series.

With $x=1$, we have

$$1+\frac{1}{\sqrt{2}}+\frac{1}{\sqrt{3}}+\frac{1}{\sqrt{4}}+\ldots+\frac{1}{\sqrt{n}}+\ldots$$

This series was shown to be divergent in § 96 (2, example 2).

Series (24) is thus seen to converge at the left-hand end of the interval of convergence and to diverge at the right-hand end; in other words, the domain of convergence is the set of x satisfying $-1\leqslant x<+1$.

Example 4. We take the series

$$1+\frac{x}{1}+\frac{x^2}{2!}+\frac{x^3}{3!}+\ldots+\frac{x^n}{n!}+\ldots \tag{25}$$

We have

$$\lim_{n\to\infty}\left|\frac{u_{n+1}(x)}{u_n(x)}\right| = \lim_{n\to\infty}\frac{|x^{n+1}|\,n!}{(n+1)!\,|x^n|}$$

$$=\lim_{n\to\infty}\left(\frac{1\cdot 2\cdot 3\ \ldots\ n}{1\cdot 2\cdot 3\ldots n\,(n+1)}\,|x|\right)=\lim_{n\to\infty}\frac{|x|}{n+1}=0;$$

hence $\lim\limits_{n\to\infty}\left|\dfrac{u_{n+1}(x)}{u_n(x)}\right|=0$ for any value of x, and series (25) is convergent throughout the number axis.

Example 5. We take the series

$$x+2!\,x^2+3!\,x^3+\ldots+n!\,x^n+\ldots$$

With $x \neq 0$ we have

$$\lim_{n\to\infty}\left|\frac{u_{n+1}(x)}{u_n(x)}\right| = \lim_{n\to\infty}\frac{(n+1)!\,|x^{n+1}|}{n!\,|x^n|} = \lim_{n\to\infty}[(n+1)\,|x|] = +\infty .$$

The series is therefore convergent only with $x=0$.

§ 101. Differentiation and integration of power series. Let the power series

$$a_0+a_1x+a_2x^2+\ldots+a_nx^n+\ldots \tag{20}$$

be convergent in the interval $(-R, R)$ $(R>0)$ and have the sum $f(x)$, a function defined in this interval.

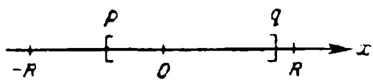


FIG. 115

It is shown in the more comprehensive courses of analysis that the series consisting of the derivatives of the terms of (20), i.e.

$$a_1+2a_2x+3a_3x^2+4a_4x^3+\ldots+na_nx^{n-1}+\ldots \tag{26}$$

has the same radius of convergence R as series (2), and that the sum of series (26) is the derivative $f'(x)$ of the sum $f(x)$ of series (20), i.e.

$$f'(x)=a_1+2a_2x+3a_3x^2+\ldots+na_nx^{n-1}+\ldots$$

This is expressed by saying that series (20) is differentiable term by term for every x in its interval of convergence.

It is also shown that series (20) can be integrated term by term over any segment $[p, q]$ lying inside its interval of convergence $(-R, R)$ (fig. 115), i.e. that

$$\int_p^q f(x)\,dx=\int_p^q a_0\,dx+\int_p^q a_1x\,dx+\int_p^q a_2x^2\,dx+\ldots$$
$$\ldots+\int_p^q a_nx^n\,dx+\ldots$$

§ 102. Expansions of the functions $\ln(1+x)$ and $\arctan x$ in power series. 1. We take the power series

$$1-x+x^2-x^3+\ldots+(-1)^n x^n+\ldots \tag{27}$$

This is a geometrical progression in $-x$ and therefore converges for $|-x|=|x|<1$, i.e. in the interval $(-1, +1)$. The sum of progression (27) is equal to $\frac{1}{1+x}$. We can thus write

$$\frac{1}{1+x}=1-x+x^2-x^3+\ldots+(-1)^n x^n+\ldots \tag{28}$$

The function $\frac{1}{1+x}$ is expressed by equation (28) as the sum of a series, or in other words, is *expanded* as a power series.

We notice that

$$\int_0^x \frac{dx}{1+x}=[\ln(1+x)]_0^x=\ln(1+x)-\ln 1=\ln(1+x).$$

Since a power series can be integrated term by term over any interval lying inside the interval of convergence, we can integrate both sides of (28) from 0 to x for any x inside $(-1, 1)$, which gives us

$$\left.\begin{aligned}\int_0^x \frac{dx}{1+x}=\int_0^x dx-\int_0^x x\,dx+\int_0^x x^2\,dx-\\ -\int_0^x x^3\,dx+\ldots+(-1)^n\int_0^x x^n\,dx+\ldots,\end{aligned}\right\}$$

or

$$\ln(1+x)=[x]_0^x-\left[\frac{x^2}{2}\right]_0^x+\left[\frac{x^3}{3}\right]_0^x-\left[\frac{x^4}{4}\right]_0^x+\ldots$$

$$\ldots+(-1)^n\left[\frac{x^{n+1}}{n+1}\right]_0^x+\ldots,$$

i.e.
$$\ln(1+x)=x-\frac{x^2}{2}+\frac{x^3}{3}-\frac{x^4}{4}+\ldots+(-1)^n\frac{x^{n+1}}{n+1}+\ldots \quad (29)$$

We have obtained the expansion of $\ln(1+x)$ as a power series. Expansion (29) is valid for $|x|<1$. It can be shown to remain valid for $x=1$, whence we have

$$\ln 2=1-\frac{1}{2}+\frac{1}{3}-\frac{1}{4}+\ldots$$

Series (29) is used to form series for evaluating logarithms.

2. Similarly, starting from the series

$$1-x^2+x^4-x^6+\ldots+(-1)^n x^{2n}+\ldots, \quad (30)$$

we can find the expansion of arc tan x as a power series. For series (30) is a geometrical progression in $-x^2$ and is therefore convergent for $|-x^2|=x^2<1$, i.e. with $|x|<1$, or in other words, with x in the interval $(-1, 1)$. The sum of progression (30) is equal to $\frac{1}{1+x^2}$ for $|x|<1$. Hence we have for $|x|<1$:

$$\frac{1}{1+x^2}=1-x^2+x^4-x^6+\ldots+(-1)^n x^{2n}+\ldots$$

On integrating this equation term by term from 0 to x (where $|x|<1$), we obtain

$$\int_0^x \frac{dx}{1+x^2}=$$

$$=\int_0^x dx-\int_0^x x^2\,dx+\int_0^x x^4\,dx-\int_0^x x^6\,dx+\ldots+(-1)^n\int_0^x x^{2n}\,dx+\ldots$$

or

$$\text{arc tan}\, x\,\Big|_0^x=[x]_0^x-\left[\frac{x^3}{3}\right]_0^x+$$

$$+\left[\frac{x^5}{5}\right]_0^x-\left[\frac{x^7}{7}\right]_0^x+\ldots+(-1)^n\left[\frac{x^{2n+1}}{2n+1}\right]_0^x+\ldots.$$

which gives us the expansion of arc tan x as a power series:

$$\operatorname{arc\,tan} x = x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \ldots + (-1)^n \frac{x^{2n+1}}{2n+1} + \ldots \tag{31}$$

It can be shown that this expansion remains valid for $x = \pm 1$. If we take $x = 1$, $\operatorname{arc\,tan} 1 = \frac{\pi}{4}$, and we obtain the series

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \ldots + (-1)^n \frac{1}{2n+1} + \ldots \tag{32}$$

This series is not suitable, however, for evaluating the number π, since a very large number of terms has to be taken to attain a reasonable accuracy (for instance, 449 terms are needed to find $\frac{\pi}{4}$ to an accuracy of 0·001). We shall show in § 104 that the number π can be evaluated by means of a series expansion of arc sin x.

These examples of expanding two functions, $\ln(1+x)$ and arc tan x, as power series suggest the possibility of expanding any function $f(x)$ in the same way. We shall discuss this question in the next two sections.

§ 103. Maclaurin series. Let us assume that the function $f(x)$ can be expanded as a power series (20), viz.

$$f(x) = a_0 + a_1 x + a_2 x^2 + \ldots + a_n x^n + \ldots, \tag{20}$$

converging in the interval $(-R, R)$, where $R > 0$.

We saw in § 101 that series (20) can be differentiated term by term in the interval $(-R, R)$, so that

$$f'(x) = a_1 + 2a_2 x + 3a_3 x^2 + \\ + 4a_4 x^3 + \ldots + na_n x^{n-1} + \ldots \tag{33}$$

Series (33) is a power series converging in the same interval as series (20) and can therefore also be differentiated term by term in the interval $(-R, R)$; the second derivative $f''(x)$ is now expanded in turn as a power series, converging in the interval $(-R, R)$, and so on. We arrive in this way at the sequence of equations:

$$\left.\begin{aligned}
f''(x) &= 1\cdot 2a_2+2\cdot 3a_3x+3\cdot 4a_4x^2+\dots\\
&\qquad +n(n-1)a_nx^{n-2}+\dots\\
f'''(x) &= 1\cdot 2\cdot 3a_3+2\cdot 3\cdot 4a_4x+\dots\\
&\qquad +n(n-1)(n-2)a_nx^{n-3}+\dots\\
f^{IV}(x) &= 1\cdot 2\cdot 3\cdot 4a_4+\dots\\
&\qquad +n(n-1)(n-2)(n-3)a_nx^{n-4}+\dots\\
&\cdot\qquad\qquad\cdot\qquad\qquad\cdot\\
f^{(n)}(x) &= 1\cdot 2\cdot 3\dots na_n+\\
&\qquad +2\cdot 3\dots n(n+1)a_{n+1}x+\dots
\end{aligned}\right\} \tag{34}$$

On now putting $x=0$ in equations (20), (33) and (34), we find that

$$\left.\begin{aligned}
&a_0=f(0),\quad a_1=f'(0),\quad a_2=\frac{f''(0)}{2!},\quad a_3=\frac{f'''(0)}{3!},\\
&a_4=\frac{f^{IV}(0)}{4!},\ \dots,\ a_n=\frac{f^{(n)}(0)}{n!},\ \dots
\end{aligned}\right\} \tag{35}$$

Substitution of these values for the coefficients $a_0, a_1, \dots, a_n, \dots$ in equation (20) gives us the formula

$$\left.\begin{aligned}
f(x)=f(0)+\frac{f'(0)}{1!}x+\frac{f''(0)}{2!}x^2+\frac{f'''(0)}{3!}x^3+\\
\frac{f^{IV}(0)}{4!}x^4+\dots+\frac{f^{(n)}(0)}{n!}x^n+\dots
\end{aligned}\right\} \tag{36}$$

The series on the right-hand side of equation (36) is known as the Maclaurin series for the function $f(x)$.

We conclude from this that, *if a function $f(x)$ can be expanded as a power series* (20) *in the interval* $(-R, R)$, *this series must be the Maclaurin series for $f(x)$.*

Furthermore, *if a function $f(x)$ is represented by a power series* (20) *in the interval* $(-R, R)$, *the function must have derivatives of all orders in this interval.*

It follows from what has been said that series (29) and (31) (§ 102) represent the Maclaurin series for the functions $\ln(1+x)$ and arc tan x respectively. This may easily be verified by working out the coefficients of the series in accordance with expressions (35); we suggest that the reader do this for the first few coefficients as an exercise.

If a function $f(x)$ has derivatives of all orders in some interval with centre at $x=0$, a Maclaurin series can always be formally composed for the function. It may happen, however, that the radius of convergence of the Maclaurin series obtained is zero, in which case equation (36) is only valid for $x=0$. In this case, equation (36) becomes valueless. Apart from this, there are functions with the following peculiarity: the Maclaurin series for the function converges inside a certain domain *yet its sum is not equal to the function* for any value of x other than $x=0$.

We can see from these remarks that the existence of derivatives of all orders is not the only condition that must be satisfied by a function in order for it to be expandable as a power series. Instead of going deeper into this question, we shall merely mention here that the functions e^x, $\sin x$, $\cos x$, $(1+x)^\alpha$ (where α is any real number differing from zero) and arc sin x can all in fact be expanded as power series (these being the Maclaurin series for the functions). In other words, each of these functions can be expressed as the sum of a power series within the domain of convergence of the latter.

§ 104. Expansion of the functions e^x, $\sin x$, $\cos x$, $(1+x)^\alpha$ and arc sin x as power series. We have just mentioned that the functions referred to in the present heading can be expanded as power series of the form (20), these expansions being in fact the Maclaurin series for the functions. We may therefore proceed as follows to obtain the actual expansions as power series:

1) work out the coefficients of the Maclaurin series for the function;
2) form the Maclaurin series;
3) find the interval of convergence of the series and consider the question of convergence at the ends of the interval.

From what has been said, the sum of the series obtained will be equal to the function in the domain of convergence, so that the series will in fact represent the expansion of the function concerned as a power series (within its domain of convergence).

1) Expansion of the function e^x as a power series. Let $f(x)=e^x$. For any x we have

$$f'(x)=e^x, \quad f''(x)=e^x, \quad f'''(x)=e^x, \dots, f^{(n)}(x)=e^x, \dots$$

Setting $x=0$, we obtain

$$f(0)=e^0=1, \quad f'(0)=1, \quad f''(0)=1,$$
$$f'''(0)=1, \dots, f^{(n)}(0)=1, \dots$$

We now evaluate the coefficients of the Maclaurin series in accordance with expressions (35) as

$$a_0=f(0)=1; \quad a_1=f'(0)=1; \quad a_2=\frac{f''(0)}{2!}=\frac{1}{2!},$$
$$a_3=\frac{f'''(0)}{3!}=\frac{1}{3!}, \dots, a_n=\frac{f^{(n)}(0)}{n!}=\frac{1}{n!}, \dots$$

Substitution of these values for the coefficients in formula gives us the Maclaurin series for e^x as

$$1+\frac{x}{1}+\frac{x^2}{2!}+\frac{x^3}{3!}+\dots+\frac{x^n}{n!}+\dots$$

We saw in example 4 of § 100 that the above series is convergent in the interval $(-\infty, +\infty)$. Hence the expansion

$$e^x=1+\frac{x}{1}+\frac{x^2}{2!}+\frac{x^3}{3!}+\dots+\frac{x^n}{n!}+\dots \tag{37}$$

is valid for any x.

If we put $x=1$ in equation (37), we arrive at the relationship

$$e=1+1+\frac{1}{2!}+\frac{1}{3!}+\dots+\frac{1}{n!}+\dots, \tag{38}$$

which gives the number e as the sum of a series.

The terms of this series are rapidly decreasing in magnitude and it is therefore suitable for evaluating e. We show that e can in fact be obtained correct to four figures by taking only the first seven terms of series (38). The error involved in neglecting the remaining terms is

$$\delta=\frac{1}{7!}+\frac{1}{8!}+\dots$$

This becomes, on taking the first term outside the brackets.

$$\delta=\frac{1}{7!}\left(1+\frac{1}{8}+\frac{1}{8.9}+\frac{1}{8.9.10}+\dots\right).$$

If we now substitute 7 for all the factors (which are >7) in the denominators in the brackets, we are clearly justified in writing

$$\delta<\frac{1}{7!}\left(1+\frac{1}{7}+\frac{1}{7^2}+\frac{1}{7^3}+\dots\right).$$

The sum of the progression inside the brackets is $\dfrac{1}{1-\dfrac{1}{7}}=\dfrac{7}{6}$.

Hence

$$\delta<\frac{1}{1.2.3.4.5.6.7}\cdot\frac{7}{6}<0{\cdot}0003\,.$$

We now write down the first seven terms of the series correct to four decimal places, as follows:

$$1=1{\cdot}0000$$

$$1=1{\cdot}0000$$

$$\frac{1}{2!}=0{\cdot}5000$$

$$\frac{1}{3!}=0{\cdot}1667$$

$$\frac{1}{4!}=0{\cdot}0417$$

$$\frac{1}{5!}=0{\cdot}0083$$

$$\frac{1}{6!}=0{\cdot}0014$$

$$\overline{\text{sum}=2{\cdot}7181}$$

The top three of these equations are accurate, whilst the error in each of the remaining four does not exceed 0·00005. Hence the error in the sum does not exceed 0·0002, and the total error does not exceed

$$0{\cdot}0002+\delta<0{\cdot}0002+0{\cdot}0003=0{\cdot}0005.$$

Hence the value 2·718 for e is correct to four significant figures.

2) Expansion of $\sin x$ as a power series.

Let $f(x)=\sin x$. We have

$$f'(x)=\cos x\,, \quad f''(x)=-\sin x\,, \quad f'''(x)=-\cos x\,, \quad f^{IV}(x)=\sin x\,,\ \ldots\,;$$

whence

$$\left.\begin{array}{l} f(0)=0\,,\ f'(0)=1\,,\ f''(0)=0\,,\ f'''(0)=-1\,,\ f^{IV}(0)=0\,,\ \ldots \\ a_0=0\,,\ a_1=1\,,\ a_2=0\,,\ a_3=-\dfrac{1}{3!}\,,\ a_4=0\,,\ a_5=\dfrac{1}{5!}\,,\ \ldots \end{array}\right|$$

Substitution of these values for the coefficients in formula (36) gives us the Maclaurin series for $\sin x$ as

$$x-\frac{x^3}{3!}+\frac{x^5}{5!}-\frac{x^7}{7!}+\ldots+(-1)^{n-1}\frac{x^{2n-1}}{(2n-1)!}+\ldots$$

This series may easily be seen to be convergent for all real x.

We thus have the expansion for any x as

$$\sin x=x-\frac{x^3}{3!}+\frac{x^5}{5!}-\frac{x^7}{7!}+\ldots+(-1)^{n-1}\frac{x^{2n-1}}{(2n-1)!}+\ldots \tag{39}$$

It can be shown that the difference between the sum of an alternating series satisfying the conditions of Leibniz's theorem (§ 97) and any given partial sum of the series has a smaller absolute

value than the absolute value of the first of the neglected terms. This gives us a simple practical means of estimating the error when using alternating series for approximations.

As an example, we shall evaluate sin 1° to an accuracy of 0·00001.

We notice first of all that the variable x in (39) is assumed to be expressed in radians. For an angle of 1°, $x=\frac{\pi}{180}=0{\cdot}017453\ldots<0{\cdot}02$. Hence even the second term $\frac{x^3}{3!}$ of expansion (39) has an absolute value less than $\frac{0{\cdot}02^3}{6}<0{\cdot}000002$. If we work out sin 1° by taking only the first term of the expansion and neglecting all the rest, the absolute value of the error involved will be less than 0·000002, showing that all the figures in the approximate value sin 1° = 0·01745 are therefore correct.

3) Expansion of cos x as a power series.

The same method can be used as when expanding sin x, i.e. we form the Maclaurin series for cos x. The same result can be achieved more simply, however, if we differentiate term by term the expansion for sin x. Differentiation of (39) gives us

$$\cos x=1-\frac{3x^2}{3!}+\frac{5x^4}{5!}-\frac{7x^6}{7!}+\ldots+(-1)^{n-1}\frac{(2n-1)\,x^{2n-2}}{(2n-1)!}+\ldots$$

or

$$\cos x=1-\frac{x^2}{2!}+\frac{x^4}{4!}-\frac{x^6}{6!}+\ldots+(-1)^{n-1}\frac{x^{2n-2}}{(2n-2)!}+\ldots \tag{39*}$$

Series (39*) is convergent, and therefore the expansion obtained for cos x is valid, for all real x.

4) The expansion of $(1+x)^\alpha$ as a power series, where α is any real number differing from zero.

Let $f(x)=(1+x)^\alpha$. We have

$$f'(x)=\alpha(1+x)^{\alpha-1}, \quad f''(x)=\alpha(\alpha-1)(1+x)^{\alpha-2},$$

$$f'''(x)=\alpha(\alpha-1)(\alpha-2)(1+x)^{\alpha-3}, \ldots, f^{(n)}(x)$$

$$=\alpha(\alpha-1)(\alpha-2)\ldots(\alpha-n+1)(1+x)^{\alpha-n}, \ldots$$

On putting $x=0$ in all these equations, we obtain

$$f(0)=1,\ f'(0)=\alpha,\ f''(0)=\alpha(\alpha-1),$$
$$f'''(0)=\alpha(\alpha-1)(\alpha-2),\ \ldots$$
$$\ldots,\ f^{(n)}(0)=\alpha(\alpha-1)(\alpha-2)\ldots(\alpha-n+1),\ \ldots$$

We use expressions (35) to evaluate the coefficients of the Maclaurin series for $(1+x)^\alpha$ as

$$\left.\begin{aligned} &a_0=1,\quad a_1=\alpha,\quad a_2=\frac{\alpha(\alpha-1)}{2!},\quad a_3=\frac{\alpha(\alpha-1)(\alpha-2)}{3!},\ \ldots,\\ &a_n=\frac{\alpha(\alpha-1)(\alpha-2)\ldots(\alpha-n+1)}{n!},\ \ldots \end{aligned}\right\} \qquad (40)$$

Thus the Maclaurin series for $(1+x)^\alpha$ is

$$1+\alpha x+\frac{\alpha(\alpha-1)}{2!}x^2+\frac{\alpha(\alpha-1)(\alpha-2)}{3!}x^3+\ldots$$
$$\ldots+\frac{\alpha(\alpha-1)(\alpha-2)\ldots(\alpha-n+1)}{n!}x^n+\ldots$$

We find its interval of convergence on the assumption that α is not a positive integer. We have

$$\lim_{n\to\infty}\frac{u_{n+1}(x)}{u_n(x)}=\lim_{n\to\infty}\frac{|\alpha(\alpha-1)(\alpha-2)\ldots(\alpha-n+1)(\alpha-n)x^{n+1}|\,n!}{(n+1)!\,|\alpha(\alpha-1)(\alpha-2)\ldots(\alpha-n+1)x^n|}$$
$$=\lim_{n\to\infty}\left(\frac{|\alpha-n|}{n+1}|x|\right)=\lim_{n\to\infty}\frac{\left|\dfrac{\alpha}{n}-1\right|}{1+\dfrac{1}{n}}|x|=|x|.$$

We conclude from this that the series is convergent for $|x|<1$, i.e. in the interval $(-1, 1)$. The expansion

$$(1+x)^\alpha=1+\alpha x+\frac{\alpha(\alpha-1)}{2!}x^2+\frac{\alpha(\alpha-1)(\alpha-2)}{3!}x^3+\ldots$$

$$\dots+\frac{\alpha(\alpha-1)(\alpha-2)\dots(\alpha-n+1)}{n!}x^n+\dots \qquad (41)$$

is therefore valid in the interval $(-1, 1)$. The sum of the series is equal to $(1+x)^\alpha$ at $x=\pm 1$ only if the series happens to be convergent for $x=\pm 1$.

Series (41) is known as the *binomial* series.

If α is a positive integer, say 3, we find on evaluating the coefficients $a_0, a_1, a_2, \dots$ in accordance with (40) that

$$a_0=1\ ; \quad a_1=3\ ; \quad a_2=\frac{3(3-1)}{2!}=3\ ; \quad a_3=\frac{3(3-1)(3-2)}{3!}=1:$$

$$a_4=\frac{3(3-1)(3-2)(3-3)}{4!}=0;$$

$$a_5=\frac{3(3-1)(3-2)(3-3)(3-4)}{5!}=0;$$

$$a_6=0 \text{ and so on.}$$

Series (41) terminates at the fourth term in this case, and the equation becomes

$$(1+x)^3=1+3x+3x^2+x^3.$$

In general, if α is a positive integer m, we have

$$a_m=\frac{m(m-1)\dots(m-m+1)}{m!}, \quad a_{m+1}=\frac{m(m-1)\dots(m-m)}{(m+1)!}=0.$$

$$a_{m+2}=0 \quad \text{and so on.}$$

Series (41) becomes a polynominal of degree m in this case, and we arrive at the equation

$$(1+x)^m=1+mx+\frac{m(m-1)}{2!}x^2+\frac{m(m-1)(m-2)}{3!}x^3+\dots$$

$$\dots+\frac{m(m-1)\dots(m-m+1)}{m!}x^m. \qquad (42)$$

This equation is sometimes called a Newton binomial. It is clearly valid for any x, and not just for $|x|<1$.

5) Expansion of arc sin x as a power series.

We have

$$\int_0^x \frac{dx}{\sqrt{1-x^2}} = [\text{arc sin } x]_0^x = \text{arc sin } x.$$

We expand the integrand $\frac{1}{\sqrt{1-x^2}} = (1-x^2)^{-\frac{1}{2}}$ as a series in accordance with expression (41) by substituting $x=-x^2$ and $\alpha = -\frac{1}{2}$. We obtain

$$\frac{1}{\sqrt{1-x^2}} = 1+\frac{1}{2}x^2+\frac{1\cdot 3}{2\cdot 4}x^4+\frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}x^6+\frac{1\cdot 3\cdot 5\cdot 7}{2\cdot 4\cdot 6\cdot 8}x^8+\ldots \quad (43)$$

Integrating both sides of this equation from 0 to x ($|x|<1$) we obtain

$$\int_0^x \frac{dx}{\sqrt{1-x^2}} = \int_0^x dx+\frac{1}{2}\int_0^x x^2\,dx+\frac{1\cdot 3}{2\cdot 4}\int_0^x x^4\,dx+\frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}\int_0^x x^6\,dx+$$
$$+\frac{1\cdot 3\cdot 5\cdot 7}{2\cdot 4\cdot 6\cdot 8}\int_0^x x^8\,dx+\ldots,$$

whence

$$\text{arc sin } x = x+\frac{1}{2}\cdot\frac{x^3}{3}+\frac{1\cdot 3}{2\cdot 4}\cdot\frac{x^5}{5}+\frac{1\cdot 3\cdot 5}{2\cdot 4\cdot 6}\cdot\frac{x^7}{7}+$$
$$+\frac{1\cdot 3\cdot 5\cdot 7}{2\cdot 4\cdot 6\cdot 8}\cdot\frac{x^9}{9}+\ldots \quad (44)$$

Since we started from expansion (43), valid only for $|-x^2| = x^2<1$, i.e. for $|x|<1$, expansion (44) is certainly valid inside the interval $(-1, +1)$. Series (44) may easily be shown to converge

for $x=\pm 1$; the sum of series (44) is therefore arc sin x throughout the segment $[-1, +1]$, and not only inside the open interval $(-1, +1)$.

We shall use expansion (44) to work out the number π to three figures. We put $x=\frac{1}{2}$, when $\arcsin \frac{1}{2}=\frac{\pi}{6}$. Hence

$$\frac{\pi}{6}=\arcsin \frac{1}{2}=\frac{1}{2}+\frac{1}{2}\cdot\frac{1}{3.2^3}+\frac{1.3}{2.4}\cdot\frac{1}{5.2^5}+\frac{1.3.5}{2.4.6}\cdot\frac{1}{7.2^7}+\dots$$

or

$$\frac{\pi}{6}=\frac{1}{2}+\frac{1}{6}\cdot\frac{1}{2^3}+\frac{3}{40}\cdot\frac{1}{2^5}+\frac{5}{112}\cdot\frac{1}{2^7}+\dots$$

or

$$\frac{\pi}{6}=\frac{1}{2}+\frac{1}{48}+\frac{3}{1280}+\frac{5}{14336}+\dots$$

We confine ourselves to the four terms written since it can be shown that the error involved is $\delta<0{\cdot}0002$, as when we evaluated the number e. We have

$$\frac{1}{2}=0{\cdot}5000$$

$$\frac{1}{48}=0{\cdot}0208$$

$$\frac{3}{1280}=0{\cdot}0023$$

$$\frac{5}{14336}=0{\cdot}0003$$

$$\text{sum}=0{\cdot}5234$$

The error in the sum is less than $0{\cdot}0002$; adding to this the error δ, the total error must be less than $0{\cdot}0004$; on rounding off, we get three correct figures for the approximate value of $\frac{\pi}{6}=0{\cdot}523$.

To find the approximate value of $\frac{\pi}{6}$ we have to multiply by 6, which gives us 3·138. On multiplying by 6 the error also increases 6 times, but remains less than 0·003. We therefore round off the value 3·138 and get $\pi \approx 3{\cdot}14$, in which all three figures are correct.

§ 105. Series in the complex plane. 1. The concept of series may be extended to include the case when the terms are complex istead of real numbers. We now have series of the form

$$(u_1+iv_1)+(u_2+iv_2)+\ldots+(u_n+iv_n)+\ldots, \tag{45}$$

where $u_1, u_2, \ldots, u_n, \ldots, v_1, v_2, \ldots, v_n, \ldots$ are real numbers and $i=\sqrt{-1}$.

Series (45) is said to be *convergent* if the series made up of the real and imaginary parts of its terms converge separately, i.e. if the series

$$u_1+u_2+\ldots+u_n+\ldots \tag{46}$$

and

$$v_1+v_2+\ldots+v_n+\ldots \tag{47}$$

are convergent.

Let U_n denote the sum of the first n terms of series (46) and V_n the sum of the first n terms of (47). When the series are convergent, the limits exist, i.e.

$$\lim_{n\to\infty} U_n = U \quad \text{and} \quad \lim_{n\to\infty} V_n = V.$$

The complex number

$$U+iV \tag{47/1}$$

is termed the sum of series (45).

Theorem. *Series (45) is convergent if the series made up of the moduli of its terms is convergent.**

* The modulus $|a+bi|$ of the complex number $a+bi$ is the real number $|a+bi|=\sqrt{a^2+b^2}$.

Proof. By hypothesis, the series

$$\sqrt{u_1^2+v_1^2}+\sqrt{u_2^2+v_2^2}+\ldots+\sqrt{u_n^2+v_n^2}+\ldots$$

is convergent. Since

$$|u_n| \leqslant \sqrt{u_n^2+v_n^2}$$

and

$$|v_n| \leqslant \sqrt{u_n^2+v_n^2}$$

($n=1, 2, \ldots$), series (46) and (47) must be convergent (absolutely) by the comparison test for series. Series (45) is now also convergent, by definition.

2. We can also consider power series in the complex plane, of she form

$$c_0+c_1 z+c_2 z^2+\ldots+c_n z^n+\ldots,$$

where

$$c_n=a_n+ib_n, \quad z=x+iy \quad (n=0, 1, 2, \ldots).$$

Here a_n and b_n are real numbers, whilst x and y are independent variables taking only real values.

By the theorem just proved, this series will be convergent if the teries formed from the moduli of its terms converges, i.e. if

$$|c_0|+|c_1||z|+|c_2||z^2|+\ldots+|c_n||z^n|+\ldots$$

is convergent.

§ **106. Euler's formula***. If x is a real variable, we know that the expansion

$$e^x=1+\frac{x}{1!}+\frac{x^2}{2!}+\frac{x^3}{3!}+\ldots \tag{37}$$

holds for any x.

---

* L. Euler (1707—1783) was one of the greatest mathematicians of all time. A Swiss by birth, he spent most of his life in Russia, and became a member of the St. Petersburg Academy of Sciences.

If $z=x+iy$, we define

$$e^z=1+\frac{z}{1}+\frac{z^2}{2!}+\frac{z^3}{3!}+\ldots \tag{48}$$

We form the series of the moduli of the terms of series (48):

$$1+\frac{\sqrt{x^2+y^2}}{2!}+\frac{(\sqrt{x^2+y^2})^2}{2!}+\ldots+\frac{(\sqrt{x^2+y^2})^n}{n!}+\ldots$$

We find on applying d'Alembert's test that

$$\lim_{n\to\infty}\frac{(\sqrt{x^2+y^2})^{n+1}\,n!}{(n+1)!\,(\sqrt{x^2+y^2})^n}=\lim_{n\to\infty}\frac{\sqrt{x^2+y^2}}{(n+1)}=0$$

for any values of x and y. Series (48) is therefore convergent for all values of the complex number z, so that the function e^z is defined for all values of z.

In particular, with $z=ix$ (x is a real variable), we have

$$e^{ix}=1+\frac{(ix)}{1!}+\frac{(ix)^2}{2!}+\frac{(ix)^3}{3!}+\ldots \tag{49}$$

Since

$$i^2=-1,\ i^3=-i,\ i^4=1,\ i^5=i,\ i^6=-1$$

and so on, expansion (49) can be written as

$$e^{ix}=1+i\frac{x}{1!}-\frac{x^2}{2!}-i\frac{x^3}{3!}+\frac{x^4}{4!}+i\frac{x^5}{5!}-\frac{x^6}{6!}-i\frac{x^7}{7!}+\frac{x^8}{8!}+\ldots$$

or

$$e^{ix}=\left(1-\frac{x^2}{2!}+\frac{x^4}{4!}-\frac{x^6}{6!}+\frac{x^8}{8!}+\ldots\right)+$$

$$+i\left(\frac{x}{1}-\frac{x^3}{3!}+\frac{x^5}{5!}-\frac{x^7}{7!}+\ldots\right).$$

By (39*) and (39), the sums of the series inside the brackets are equal to $\cos x$ and $\sin x$ respectively. We thus arrive at the following remarkable expression

$$e^{ix}=\cos x+i\sin x. \tag{50}$$

On replacing x by $-x$, we obtain

$$e^{-ix}=\cos x-i\sin x, \tag{50*}$$

Expressions (50) and (50*) are known as Euler's formulae.

Euler's formulae readily lead to the following expressions for $\cos x$ and $\sin x$ in terms of the exponential functions:

$$\cos x=\frac{e^{ix}+e^{-ix}}{2}$$

and

$$\sin x=\frac{e^{ix}-e^{-ix}}{2i}.$$

In the general case when $z=x+iy$, it can be shown that

$$e^{x+iy}=e^{x}\cdot e^{iy},$$

or by (50)

$$e^{x+iy}=e^{x}(\cos y+i\sin y).$$

As a particular result we have

$$e^{1+\frac{\pi}{2}i}=e\left(\cos\frac{\pi}{2}+i\sin\frac{\pi}{2}\right)=ei.$$

§ 107. Trigonometric series. *Periodic* processes are often encountered in science and engineering, i.e. processes which repeat themselves after a definite period of time T. To take an example, when a steam engine is running steadily its moving parts pass through their initial positions with the same velocities and accelerations after definite periods of time. An alternating current is another example of a periodic quantity. Periodic phenomena are also observed in nature, such as the tides, seasonal variations in the density of sea water, etc. The mathematical investigation of such processes led to the concept of periodic function.

If the value of a function $f(t)$ is unchanged on adding the number T to its argument, i.e.

$$f(t+T)=f(t),$$

the function is said to be *periodic with period T*.

The various quantities in the processes mentioned above are periodic functions of time t: after the period T has elapsed the periodic function $f(t)$ again takes its initial value. It is self-evident that the values of $f(t)$ start repeating themselves after each of the intervals $2T$, $3T$, ..., so that $2T$, $3T$, ... are also periods of $f(t)$; it may be mentioned, for the sake of accuracy, that T is taken to mean the least of the periods of $f(t)$.

The simplest periodic functions are $A \sin \omega t$ and $A \cos \omega t$, where A and ω are definite numbers. Let the least period of $A \sin \omega t$ be T; then

$$A \sin \omega (t+T) = A \sin \omega t$$

or

$$\sin (\omega t + \omega T) = \sin \omega t.$$

On the other hand we know from trigonometry that

$$\sin (\omega t + 2\pi) = \sin \omega t,$$

where 2π is the least of the periods of $\sin \omega t$. Comparison of the last two equations shows that $\omega T = 2\pi$, whence

$$\omega = \frac{2\pi}{T}.$$

The quantity $\frac{\omega}{2\pi}$ is called the *frequency* of the quantity $y = A \sin \omega t$. As t varies, the value of $\sin \omega t$ oscillates between -1 and $+1$; hence y oscillates between $-A$ and $+A$; the constant A is termed the *amplitude of the sinusoidal magnitude* $y = A \sin \omega t$. The graphs of sinusoidal quantities consist of wave-shaped curves; fig. 116 illustrates the graph of $y = A \sin \omega t$ for the same amplitude $A = 1$ and with $\omega = 1$, $\omega = 2$ and $\omega = \frac{1}{2}$.

Periodic processes very rarely occur in nature according to the elementary sinusoidal law $y = A \sin \omega t$ (or $y = A \cos \omega t$). We then try to represent the function expressing the process as the sum of elementary periodic functions with frequencies equal to multiplies

of $\frac{\omega}{2\pi}$ (i.e. frequencies $\frac{\omega}{2\pi}$, $\frac{2\omega}{\pi}$, $\frac{3\omega}{2\pi}$ and so on) and suitably

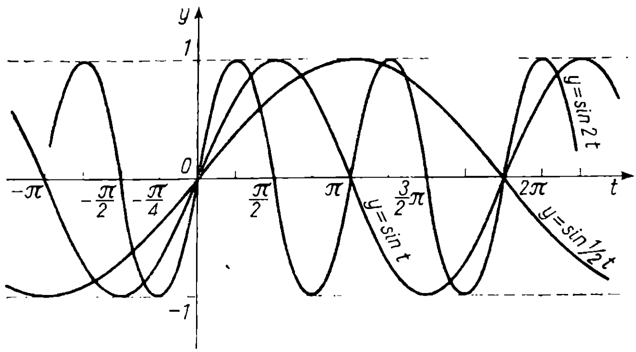


FIG. 116

chosen amplitudes. It happens that we can express *accurately* the majority of periodic functions $f(t)$ encountered in practice as the sum of a series of the form

$$f(t)=\frac{a_0}{2}+(a_1 \cos \omega t+b_1 \sin \omega t)+(a_2 \cos 2\omega t+b_2 \sin 2\omega t)+$$

$$+(a_3 \cos 3\omega t+b_3 \sin 3\omega t)+\dots$$

$$=\frac{a_0}{2}+\sum_{n=1}^{\infty} (a_n \cos n\,\omega t+b_n \sin n\,\omega t), \tag{T}$$

where a_0, a_1, b_1, a_2, b_2, ... are constants.

A series of this type is termed a *trigonometric series*, and equation (T) is said to be the expansion of $f(t)$ as a trigonometric series.

We put $\omega t=x$; then series (T) becomes

$$\frac{a_0}{2}+(a_1 \cos x+b_1 \sin x)+(a_2 \cos 2x+b_2 \sin 2x)+$$

$$+(a_3 \cos 3x+b_3 \sin 3x)+\ldots$$

$$=\frac{a_0}{2}+\sum_{n=1}^{\infty}(a_n \cos nx+b_n \sin nx). \tag{F}$$

All the terms of series (F) are periodic functions with the common period 2π*; hence the sum of series (F) has the same periodicity.

It often becomes important in applied mathematics to expand a *non-periodic* function as a trigonometric series. In view of the periodicity of the sum of series (F), it would be absurd to seek an expansion of a non-periodic function in an interval of length greater than 2π. We shall always take $[0, 2\pi]$ as our interval of length 2π in future.

§ 108. Fourier coefficients. Fourier series. 1. Suppose we know that $f(x)$ can be expanded as a trigonometric series in the interval $[0, 2\pi]$, although we do not know the actual expansion. To obtain this expansion, we have to know how to find the coefficients $a_0, a_1, b_1, a_2, b_2, \ldots$ from our knowledge of $f(x)$.

The determination of the coefficients is based on the evaluation of the integrals from 0 to 2π of the following functions: $\sin nx$, $\cos nx$, $\sin nx \sin mx$, $\cos nx \cos mx$, $\sin nx \cos mx$, where n and m are positive integers.

1) Evaluation of $\int_0^{2\pi} \sin nx\, dx$ and $\int_0^{2\pi} \cos nx\, dx$. We notice before evaluating the integrals that, if k is a positive or negative integer ($k \neq 0$), then

$$\int_0^{2\pi} \sin kx\, dx=0 \quad \text{and} \quad \int_0^{2\pi} \cos kx\, dx=0.$$

For, on substituting $kx=u$, we have $x=\frac{u}{k}$, $dx=\frac{du}{k}$, and

* Naturally 2π is not the least period for all the terms.

$$\int \sin kx\, dx = \frac{1}{k}\int \sin u\, du$$

$$= \frac{1}{k}(-\cos u) + C = -\frac{1}{k}\cos kx + C.$$

Consequently,

$$\int_0^{2\pi} \sin kx\, dx = -\frac{1}{k}[\cos kx]_0^{2\pi}$$

$$= -\frac{1}{k}(\cos 2k\pi - \cos 0) = -\frac{1}{k}(1-1) = 0.$$

We obtain similarly

$$\int_0^{2\pi} \cos kx\, dx = \frac{1}{k}[\sin kx]_0^{2\pi}$$

$$= \frac{1}{k}(\sin 2k\pi - \sin 0) = \frac{1}{k}(0-0) = 0.$$

We see at once from this that

$$1)\ \int_0^{2\pi} \sin nx\, dx = 0 \quad \text{and} \quad 2)\ \int_0^{2\pi} \cos nx\, dx = 0.$$

2) Evaluation of $\int_0^{2\pi} \sin mx \sin nx\, dx$.

We first take the case when $m \neq n$. We use the trigonometric formula $\cos(m-n)x - \cos(m+n)x = 2\sin mx \sin nx$.
This gives us

$$\int_0^{2\pi} \sin mx \sin nx\, dx$$

$$= \frac{1}{2}\int_0^{2\pi} [\cos(m-n)x - \cos(m+n)x]\, dx =$$

$$= \frac{1}{2}\left[\int_0^{2\pi} \cos(m-n)\, dx - \int_0^{2\pi} \cos(m+n)x\, dx\right].$$

Here $m-n$ and $m+n$ are integers, whilst $m+n\neq 0$ and (since $m\neq n$) $m-n\neq 0$. It follows that each of the last integrals vanishes, so that we also have

$$\int_0^{2\pi} \sin mx \sin nx\, dx = 0.$$

If $m=n$, we have

$$\int_0^{2\pi} \sin mx \sin nx\, dx = \int_0^{2\pi} \sin^2 nx\, dx$$

$$=\frac{1}{2}\int_0^{2\pi} (1-\cos 2nx)\, dx = \frac{1}{2}\left\{[x]_0^{2\pi} - \frac{1}{2n}[\sin 2nx]_0^{2\pi}\right\} = \pi.$$

3) Evaluation of $\int_0^{2\pi} \cos mx \cos nx\, dx$.

With $m\neq n$, we use the formula

$$\cos(m+n)x + \cos(m-n)x = 2\cos mx \cos nx.$$

which gives us

$$\int_0^{2\pi} \cos mx \cos nx\, dx =$$

$$=\frac{1}{2}\left[\int_0^{2\pi} \cos(m+n)x\, dx + \int_0^{2\pi} \cos(m-n)x\, dx\right].$$

Each of the integrals in the square brackets vanishes, since $m+n$ and $m-n$ are integers.

Hence $(m\neq n)$,

$$\int_0^{2\pi} \cos mx \cos nx\, dx = 0.$$

If $m=n$, we have

$$\int_0^{2\pi} \cos mx \cos nx\, dx = \int_0^{2\pi} \cos^2 nx\, dx = \frac{1}{2}\int_0^{2\pi} (1+\cos 2nx)\, dx$$

$$= \frac{1}{2}\left\{[x]_0^{2\pi} + \frac{1}{2n}[\sin 2nx]_0^{2\pi}\right\} = \pi.$$

4) Evaluation of $\int_0^{2\pi} \sin nx \cos mx\, dx$.

With $m \neq n$, we use the formula

$$\sin (m+n)\, x + \sin (m-n)\, x = 2 \sin mx \cos nx,$$

and find, precisely as above, that

$$\int_0^{2\pi} \sin mx \cos nx\, dx = 0.$$

If $m=n$, we have

$$\int_0^{2\pi} \sin mx \cos nx\, dx = \int_0^{2\pi} \sin nx \cos nx\, dx$$

$$= \frac{1}{2}\int_0^{2\pi} \sin 2nx\, dx = -\frac{1}{2n}[\cos 2nx]_0^{2\pi} = 0.$$

Thus $\int_0^{2\pi} \sin mx \cos nx\, dx = 0$ both for $m \neq n$ and $m=n$.

2. We now turn to the question of finding the coefficients of the expansion in the interval $[0, 2\pi]$ of

$$f(x) = \frac{a_0}{2} + a_1 \cos x + b_1 \sin x + a_2 \cos 2x + b_2 \sin 2x + \ldots \quad (F)$$

On integrating this equation term by term from 0 to 2π we have

$$\int_0^{2\pi} f(x)\, dx = \frac{a_0}{2}\int_0^{2\pi} dx + a_1 \int_0^{2\pi} \cos x\, dx + b_1 \int_0^{2\pi} \sin x\, dx +$$

$$+ a_2 \int_0^{2\pi} \cos 2x\, dx + b_2 \int_0^{2\pi} \sin 2x\, dx + \ldots$$

All the integrals on the right-hand side except the first are of types (1) or (2) considered above, and since we have seen that these vanish, the last equation reduces to

$$\int_0^{2\pi} f(x)\,dx=\frac{a_0}{2}\int_0^{2\pi} dx,$$

whence

$$\int_0^{2\pi} f(x)\,dx=\frac{a_0}{2}\,[x]_0^{2\pi}=a_0\pi,$$

which gives us

$$a_0=\frac{1}{\pi}\int_0^{2\pi} f(x)\,dx. \tag{51}$$

To find the coefficient a_n, where $n\neq 0$, we multiply both sides of equation (F) by $\cos nx$ then integrate term by term between the limits 0 and 2π. We have

$$\begin{aligned}\int_0^{2\pi} f(x)\cos nx\,dx&=\frac{a_0}{2}\int_0^{2\pi}\cos nx\,dx+a_1\int_0^{2\pi}\cos x\cos nx\,dx+\\&+b_1\int_0^{2\pi}\sin x\cos nx\,dx+a_2\int_0^{2\pi}\cos 2x\cos nx\,dx+\\&+b_2\int_0^{2\pi}\sin 2x\cos nx\,dx+\ldots+a_n\int_0^{2\pi}\cos nx\cos nx\,dx+\\&+b_n\int_0^{2\pi}\sin nx\cos nx\,dx+a_{n+1}\int_0^{2\pi}\cos(n+1)x\cos nx\,dx+\ldots\end{aligned} \tag{52}$$

We have seen that $\int_0^{2\pi}\cos nx\,dx=0$, $\int_0^{2\pi}\cos mx\cos nx\,dx=0$ and $\int_0^{2\pi}\sin mx\cos nx\,dx=0$ when $m\neq n$, whilst $\int_0^{2\pi}\cos nx\cos nx\,dx$ $=\int_0^{2\pi}\cos^2 nx\,dx=\pi$. All the terms of series (52) therefore vanish except $a_n\int_0^{2\pi}\cos nx\cos nx\,dx=a_n\pi$.

It follows from (52) that

$$\int_0^{2\pi} f(x)\cos nx\,dx = a_n\pi,$$

whence we obtain for the coefficient a_n

$$a_n = \frac{1}{\pi}\int_0^{2\pi} f(x)\cos nx\,dx. \qquad (53)$$

We see on comparing formulae (51) and (53) that the latter reduces to the former on putting $n=0$; in other words, all the coefficients $a_0, a_1, a_2, \ldots$ can be defined by expression (53). The first term of series (F) is written as $\frac{1}{2}a_0$ instead of simply a_0 precisely so that expression (53) can be used to define the first term as well as the second, third etc. terms.

We find the coefficients $b_n (n=1, 2, 3, \ldots)$ by multiplying both sides of equation (F) by sin nx then integrating term by term between the limits 0 and 2π. This gives us

$$\int_0^{2\pi} f(x)\sin nx\,dx = \frac{a_0}{2}\int_0^{2\pi}\sin nx\,dx + a_1\int_0^{2\pi}\cos x\sin nx\,dx +$$

$$+ b_1\int_0^{2\pi}\sin x\sin nx\,dx + \ldots + a_n\int_0^{2\pi}\cos nx\sin nx\,dx +$$

$$+ b_n\int_0^{2\pi}\sin nx\sin nx\,dx + \ldots$$

All the integrals on the right-hand side vanish except

$$\int_0^{2\pi}\sin nx\sin nx\,dx = \int_0^{2\pi}\sin^2 nx\,dx = \pi,$$

so that we have

$$\int_0^{2\pi} f(x)\sin nx\,dx = b_n\pi,$$

whence

$$b_n = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin nx \, dx \qquad (n = 1, 2, 3, \ldots). \tag{54}$$

The formulae for defining the coefficients a_n and b_n are thus:

$$a_n = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos nx \, dx \qquad (n = 0, 1, 2, 3, \ldots). \tag{53}$$

$$b_n = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin nx \, dx \qquad (n = 1, 2, 3, \ldots). \tag{54}$$

The coefficients defined by these formulae are generally known as *Fourier coefficients* * (although formulae (53) and (54) were first discovered by Euler), whilst the trigonometric series (F) constructed with the aid of the formulae is known as a *Fourier series.*

Note. Comprehensive courses of mathematical analysis prove the validity of the operations used in deducing formulae (53) and (54). The proof is extremely complicated and lies outside the scope of the present book.

3. Before turning to the applications of Fourier expansions we shall mention a rule of integration which often proves useful in evaluating Fourier coefficients. This rule, known as "integration by parts", is based on the inversion of the rule for differentiating the product of two functions with the argument x; this latter is given by the familiar formula

$$(uv)' = u'v + uv'.$$

Hence

$$\int (uv)' \, dx = \int (u'v + uv') \, dx$$

or

$$uv + C = \int v (u' \, dx) + \int u (v' \, dx).$$

* J. Fourier (1768–1830) was a celebrated French mathematician.

Noticing that $u'dx = du$ and $v'dx = dv$, we obtain

$$\int u\,dv = uv - \int v\,du, \tag{55}$$

the arbitrary constant C on the left-hand side being assumed to be included in $\int v\,du$.

Formula (55) is called the formula for integration by parts.

The formula reduces integration of the expression $u\,dv = uv'\,dx$ to integration of $v\,du = vu'\,dx$.

Suppose, for instance, that we want to find $\int x\cos x\,dx$. We put

$$x = u\,, \quad dv = \cos x\,dx,$$

so that $du = dx$, $v = \sin x$.

By formula (55) we now obtain

$$\int x\cos x\,dx = x\sin x - \int \sin x\,dx = x\sin x + \cos x + C.$$

The rule for integration by parts has enabled us to reduce the integral of the complicated integrand $x\cos x$ to the integral of the simple function $\sin x$.

There is also a rule for integration by parts for definite integrals.

We recall that

$$\int_a^b f(x)\,dx = F(b) - F(a),$$

where $F(x)$ is *any* primitive of $f(x)$. We can therefore write

$$\int_a^b f(x)\,dx = [\int f(x)\,dx]_a^b.$$

We now obtain

$$\int_a^b u\,dv = [\int u\,dv]_a^b = [uv - \int v\,du]_a^b$$

or

$$\int_a^b u\,dv = [uv]_a^b - \int_a^b v\,du. \tag{55*}$$

Suppose we want to evaluate

$$\int_0^{2\pi} x \cos 3x\, dx.$$

We put $x=u$, $\cos 3x\, dx=dv$, whence $du=dx$ and $v=\frac{1}{3}\sin 3x$.

Formula (55*) now gives us

$$\int_0^{2\pi} x \cos 3x\, dx=\frac{1}{3}[x \sin 3x]_0^{2\pi}-\frac{1}{3}\int_0^{2\pi} \sin 3x\, dx$$

$$=\frac{1}{3}(2\pi \sin 6\pi-0 \sin 0)+\frac{1}{9}[\cos 3x]_0^{2\pi}$$

$$=\frac{1}{9}(\cos 6\pi-\cos 0)=0.$$

4. We consider the function $y=f(x)$ defined in the interval $[0, 2\pi]$ as follows:

$$f(x)=\begin{cases} 0 & \text{for} \quad 0 \leqq x \leqq \pi, \\ x & \text{for} \quad \pi < x \leqq 2\pi. \end{cases}$$

By hypothesis, $f(x)=x$ for $x>\pi$ and $x\leqq 2\pi$, while $f(\pi)=0$. At the same time, $\lim\limits_{x\to\pi} f(x)=\pi$ when $x\to\pi$ while remaining *greater*

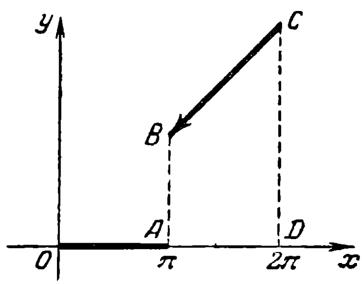


FIG. 117

than π (when x tends to π *from the right*). The function $f(x)$ thus changes by a jump, i.e. has a discontinuity when x passes through π. The graph of the function is shown in fig. 117; the arrow at

the left-hand end of the straight line representing $f(x)$ for $\pi < x \leqq 2\pi$ indicates that the point where the arrow ends, i.e. the point $B(\pi, \pi)$ does not belong to the graph (since $f(\pi)=0$ and not π!).

We have so far been concerned exclusively with continuous functions in the integral calculus. The concept of definite integral can in fact be extended to certain cases when the function has discontinuities in the interval of integration. We shall indicate how this is done by reference to the above function $f(x)$.

We divide the interval $[0, 2\pi]$ into $[0, \pi]$ and $[\pi, 2\pi]$ and write by way of definition

$$\int_0^{2\pi} f(x)\,dx = \int_0^{\pi} f(x)\,dx + \int_{\pi}^{2\pi} x\,dx.$$

It will be seen that the second integral on the right is of the function $\varphi(x)=x$, which has the value $\varphi(\pi)=\pi$ at the left-hand end of the interval $[\pi, 2\pi]$, i.e. has a value equal to the limit $\lim\limits_{x\to\pi} f(x)$ as $x\to\pi$ from the right instead of the value $f(\pi)=0$. Observing that $f(x)=0$ in the interval $[0, \pi]$, we have

$$\int_0^{2\pi} f(x)\,dx = \int_0^{\pi} 0\cdot dx + \int_{\pi}^{2\pi} x\,dx. \qquad (55^*/4)$$

Since $\int 0.\,dx = C = \text{constant}$ (because $C'=0$), we have $\int\limits_0^{\pi} 0.\,dx = [C]_0^{\pi} = C - C = 0$; also $\int\limits_{\pi}^{2\pi} x\,dx = \left[\frac{1}{2}x^2\right]_{\pi}^{2\pi} = \frac{3}{2}\pi^2$. Consequently $\int\limits_0^{2\pi} f(x)\,dx = \frac{3}{2}\pi^2$.

In introducing the function $\varphi(x)=x$ we have replaced the function $f(x)$ by two continuous functions: $y=0$ (in the interval $[0, \pi]$) and $y=x$ (in the interval $[\pi, 2\pi]$).

This method of defining $\int\limits_0^{2\pi} f(x)\,dx$ is fully consistent with the geometrical interpretation of a definite integral, for the introduction of the function $\varphi(x)=x$ in the interval $[\pi, 2\pi]$ adds the missing point $B(\pi, \pi)$ to the graph of $f(x)$ in this interval, and $\int\limits_0^{2\pi} f(x)\,dx$

now defines the area of the figure formed by the length OA and the trapezium $ABCD$ (fig. 118).

5. We shall now consider two examples of forming the Fourier series for a given function.

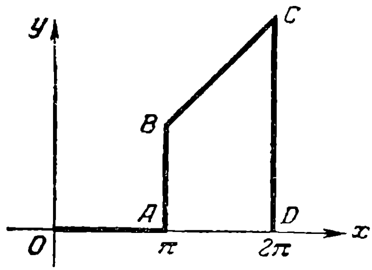


FIG. 118

Example 1. We form the Fourier series for the function discussed in the previous section, i.e. for $f(x)$ defined as follows in the interval $[0, 2\pi]$:

$$f(x) = \begin{cases} 0 \text{ for } 0 \leqq x \leqq \pi, \\ x \text{ for } \pi < x \leqq 2\pi. \end{cases}$$

We evaluate the Fourier coefficients in accordance with formulae 53) and (54). Thus

$$\frac{a_0}{2} = \frac{1}{2\pi}\int_0^{2\pi} f(x)\,dx = \frac{1}{2\pi}\left\{\int_0^{\pi} 0 \cdot dx + \int_{\pi}^{2\pi} x\,dx\right\} = \frac{1}{2\pi}\left[\frac{x^2}{2}\right]_{\pi}^{2\pi}$$

$$= \frac{1}{2\pi}\left(2\pi^2 - \frac{\pi^2}{2}\right) = \frac{3}{4}\pi.$$

We obtain with $n = 1, 2, \ldots,$

$$a_n = \frac{1}{\pi}\int_0^{2\pi} f(x)\,dx = \frac{1}{\pi}\left\{\int_0^{\pi} 0 \cdot dx + \int_{\pi}^{2\pi} x \cos nx\,dx\right\}$$

$$= \frac{1}{\pi}\int_{\pi}^{2\pi} x \cos nx\,dx.$$

We evaluate $\int\limits_{\pi}^{2\pi} x \cos nx\, dx$ by the rule for integration by parts by writing

$$\begin{array}{l|l} x = u\,, & du = dx, \\ \cos nx\, dx = dv\,. & v = \dfrac{1}{n} \sin nx. \end{array}$$

Hence,

$$\int\limits_{\pi}^{2\pi} x \cos nx\, dx = \frac{1}{n} [x \sin nx]_{\pi}^{2\pi} - \frac{1}{n} \int\limits_{\pi}^{2\pi} \sin nx\, dx$$

$$= \frac{1}{n} (2\pi \sin 2n\pi - \pi \sin n\pi) + \frac{1}{n^2} [\cos nx]_{\pi}^{2\pi}$$

$$= \frac{1}{n^2} (\cos 2\pi n - \cos \pi n)$$

and consequently,

$$a_n = \frac{1}{\pi n^2} (\cos\ 2\pi n - \cos\ \pi n).$$

We thus have:

$$a_1 = \frac{1}{\pi} (\cos 2\pi - \cos \pi) = \frac{2}{\pi},$$

$$a_2 = \frac{1}{2^2 \pi} (\cos 4\pi - \cos 2\pi) = 0,$$

$$a_3 = \frac{1}{3^2 \pi} (\cos 6\pi - \cos 3\pi) = \frac{2}{3^2 \pi},$$

$$a_4 = \frac{1}{4^2 \pi} (\cos 8\pi - \cos 4\pi) = 0,$$

In general, $a_n = 0$ with n even, whilst with n odd, $n = 2k - 1$ $(k = 1, 2, \ldots)$, we have

$$a_{2k-1}=\frac{1}{(2k-1)^2\pi}\,[\cos 2\,(2k-1)\,\pi-\cos\,(2k-1)\,\pi]$$

$$=\frac{2}{(2k-1)^2\pi}\,.$$

Further,

$$b_n=\frac{1}{\pi}\int_0^{2\pi}f(x)\sin nx\,dx=\frac{1}{\pi}\left\{\int_0^{\pi}0\cdot dx+\int_{\pi}^{2\pi}x\sin nx\,dx\right\}$$

$$=\frac{1}{\pi}\int_{\pi}^{2\pi}x\sin nx\,dx$$

$$=\frac{1}{\pi}\left\{-\frac{1}{n}\,[x\cos nx]_{\pi}^{2\pi}+\frac{1}{n}\int_{\pi}^{2\pi}\cos nx\,dx\right\}$$

$$=\frac{1}{\pi}\left\{-\frac{1}{n}\,(2\pi\cos 2\pi n-\pi\cos\pi n)+\frac{1}{n^2}\,[\sin nx]_0^{2\pi}\right\}$$

$$=-\frac{1}{n}\,(2\cos 2\pi n-\cos\pi n).$$

Hence

$$b_1=-(2\cos 2\pi-\cos\pi)=-3,$$

$$b_2=-\frac{1}{2}\,(2\cos 4\pi-\cos 2\pi)=-\frac{1}{2},$$

$$b_3=-\frac{1}{3}\,(2\cos 6\pi-\cos 3\pi)=-\frac{3}{3},$$

$$b_4=-\frac{1}{4}\,(2\cos 8\pi-\cos 4\pi)=-\frac{1}{4},$$

$$b_5=-\frac{1}{5}\,(2\cos 10\pi-\cos 5\pi)=-\frac{3}{5},$$

$$b_6=-\frac{1}{6}\,(2\cos 12\pi-\cos 6\pi)=-\frac{1}{6},$$

In general, if n is odd, $n=2k-1$, we have $b_{2k-1}=-\frac{3}{2k-1}$, whilst with n even, $n=2k$, $b_{2k}=-\frac{1}{2k}$.

The Fourier series for the function in question is therefore

$$\frac{3}{4}\pi+\frac{2}{\pi}\cos x-3\sin x-\frac{1}{2}\sin 2x+\frac{2}{3^2\pi}\cos 3x-\frac{3}{3}\sin 3x-$$
$$-\frac{1}{4}\sin 4x+\frac{2}{5^2\pi}\cos 5x-\frac{3}{5}\sin 5x-\frac{1}{6}\sin 6x+\ldots$$

or

$$\frac{3}{4}\pi+\frac{2}{\pi}\left(\cos x+\frac{\cos 3x}{3^2}+\frac{\cos 5x}{5^2}+\frac{\cos 7x}{7^2}+\ldots\right)-$$
$$-\left(\frac{3}{1}\sin x+\frac{1}{2}\sin 2x+\frac{3}{3}\sin 3x+\frac{1}{4}\sin 4x+\right.$$
$$\left.+\frac{3}{5}\sin 5x+\frac{1}{6}\sin 6x+\ldots\right).$$

Example 2. We form the Fourier series for the function $f(x)=x$. We have

$$\frac{a_0}{2}=\frac{1}{2\pi}\int_0^{2\pi}x\,dx=\pi,$$

$$a_n=\frac{1}{\pi}\int_0^{2\pi}x\cos nx\,dx$$
$$=\frac{1}{\pi}\left\{\frac{1}{n}\left[x\sin nx\right]_0^{2\pi}-\frac{1}{n}\int_0^{2\pi}\sin nx\,dx\right\}=0,$$

$$b_n=\frac{1}{\pi}\int_0^{2\pi}x\sin nx\,dx$$

$$=\frac{1}{\pi}\left\{-\frac{1}{n}\left[x\cos nx\right]_0^{2\pi}+\frac{1}{n}\int_0^{2\pi}\cos nx\,dx\right\}$$

$$=-\frac{1}{\pi n}\,2\pi\cos 2\pi n=-\frac{2}{n}.$$

The required series is therefore

$$\pi-2\left(\frac{\sin x}{1}+\frac{\sin 2x}{2}+\frac{\sin 3x}{3}+\frac{\sin 4x}{4}+\ldots\right). \tag{56}$$

6. If we set $x=0$ and then $x=2\pi$ in Fourier series (56) for the function $f(x)=x$, all the terms in the brackets vanish and the sum of the series becomes π, whereas $f(0)=0$ and $f(2\pi)=2\pi$.

The sum of the series is thus seen to differ from the value of $f(x)=x$ at $x=0$ and $x=2\pi$, so that $f(x)=x$ cannot be represented by its Fourier series for these values of x.

There are certain functions whose Fourier series are divergent. In other cases, the Fourier series is convergent but its sum is not equal to the function at every point of the interval.

We have so far only established that, if a function $f(x)$ can be expanded as a trigonometric series in the interval $[0\,,2\pi]$, this series is its Fourier series, i.e. the coefficients are expressed in terms of the given function by formulae (53) and (54).

The fundamental problem in the theory of trigonometric series is clearly that of establishing the conditions that have to be satisfied for a given function to be equal to the sum of its Fourier series. A great deal of attention has been given to this problem. The theorem which we state below, without proof, is one of the simplest relating to the Fourier expansion of functions. At the same time, in spite of its simplicity, Fourier series can be applied to the investigation of a wide range of natural phenomena by means of the theorem.

THEOREM. *Let the function $f(x)$ be continuous, except possibly at a finite number of points, in the interval* $[0\,,2\pi]$, *and let it have no extrema, or only a finite number of such, in the interval; then the Fourier expansion for $f(x)$ in the interval* $[0\,,2\pi]$ *will have the sum $S(x)$ which is equal to $f(x)$ at any point where $f(x)$ is continuous.*

while at any point where $f(x)$ is discontinuous $S(x)$ is equal to the ordinate of the mid-point of the jump that occurs in the graph of $f(x)$; at the ends of the interval the values of $S(x)$ are both equal to the arithmetic mean of the values of $f(x)$ at $x=0$ and $x=2\pi$, i.e.

$$S(0)=S(2\pi)=\frac{f(0)+f(2\pi)}{2}.$$

Since the function $f(x)=x$ is continuous and has no extrema in the interval $[0, 2\pi]$, it satisfies the conditions of the last theorem. It follows from the theorem that the sum $S(x)$ of Fourier series (56) coincides with the function x at every interior point of $[0, 2\pi]$, i.e. in the open interval $(0, 2\pi)$. Moreover the theorem states that

$$S(0)=S(2\pi)=\frac{f(0)+f(2\pi)}{2}=\frac{0+2\pi}{2}=\pi.$$

We proved above, by direct evaluation of the Fourier series of $f(x)=x$ at $x=0$ and $x=2\pi$, that its sum was equal to π at both these points.

It follows from the above theorem that the function $f(x)=x$ can be expanded as a Fourier series in the open interval $(0, 2\pi)$, i.e. the following equation is valid for $0<x<2\pi$,

$$x=\pi-2\left(\frac{\sin x}{1}+\frac{\sin 2x}{2}+\frac{\sin 3x}{3}+\frac{\sin 4x}{4}+\ldots\right). \tag{57}$$

We proved in § 107 that the sum $S(x)$ of the Fourier series of a given function $f(x)$ is a periodic function. This means that the graph of the sum $S(x)$ of series (57) consists of a set of parallel straight lines (fig. 119) and of isolated points with co-ordinates $\ldots (-2\pi, \pi), (0, \pi), (2\pi, \pi), (4\pi, \pi), \ldots$

The dotted line shows the graph of $f(x)=x$ outside the interval $(0, 2\pi)$. It will be seen from this that the function itself has nothing in common with the sum of its Fourier series outside the interval $(0, 2\pi)$. This supports in a visual way our assertion in § 107 that it would be meaningless to seek the expansion of a non-periodic function as a trigonometric series in an interval of length greater than 2π.

The function

$$f(x)=\begin{cases}0 & \text{for} \quad 0 \leqq x \leqq \pi, \\ x & \text{for} \quad \pi < x \leqq 2\pi\end{cases}$$

also satisfies the conditions of the above theorem. The sum $S(x)$ of its Fourier series, found in example 1 (5.), coincides with $f(x)$

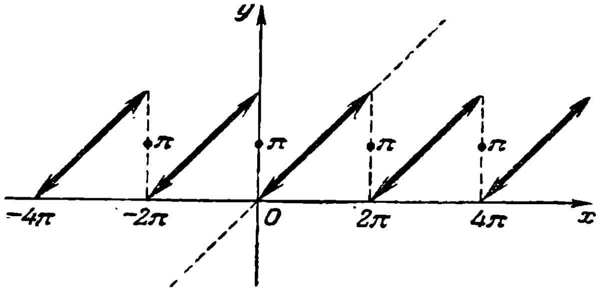


FIG. 119

at points where $f(x)$ is continuous; at the point $x=\pi$ where $f(x)$ umps, $S(\pi)=\frac{0+\pi}{2}=\frac{1}{2}\pi$; also, $S(0)=S(2\pi)=\frac{0+2\pi}{2}=\pi$.

§ 109. Fourier expansions of functions in sines only or cosines only. 1. We shall consider a function in the interval $[0, 2\pi]$ such as that illustrated in fig. 120, with the property of being symmetrical about the point $(\pi, 0)$.

It follows at once from geometrical considerations that

$$\int_0^{2\pi} f(x)\, dx=0.$$

Also, we shall prove that

$$\int_0^{2\pi} f(x) \cos nx\, dx=0.$$

This integral is the limit of the sum

$$\sum_0^{2\pi} f(x) \cos nx\, \Delta x \tag{J}$$

on the assumption that all the lengths Δx of the sub-intervals into which $[0, 2\pi]$ is arbitrarily divided tend simultaneously to zero. Since the limit of the sum (J) is the same whatever the method of subdivision of the interval $[0, 2\pi]$, we can choose the particular method that suits our purposes.

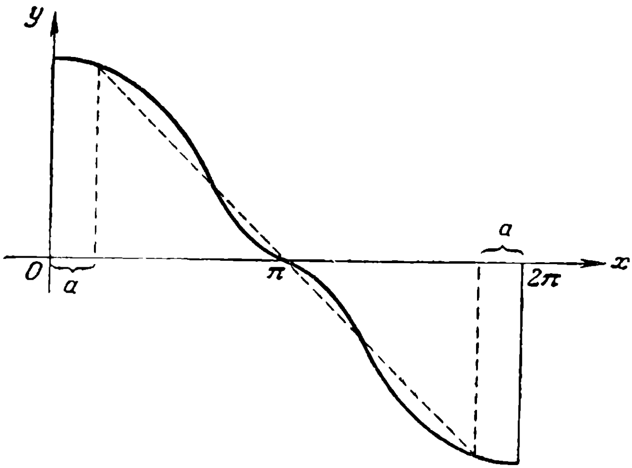


FIG. 120

Each time we take an arbitrary point $x=a$ in the interval $[0, \pi]$, we take the point $x=2\pi-a$ in the interval $[\pi, 2\pi]$. Hence our sum (J) will be made up of pairs of terms $f(a)\cos na$ and $f(2\pi-a)$ $\cos n(2\pi-a)$.

In view of the symmetry of the graph of the function with respect to the point $(\pi, 0)$, we have

$$f(2\pi-a)=-f(a).$$

Since

$$\cos n(2\pi-a)=\cos(2\pi n-na)=\cos na$$

we have

$$f(2\pi-a)\cos n(2\pi-a)=-f(a)\cos na,$$

so that

$$f(a)\cos na+f(2\pi-a)\cos n(2\pi-a)$$

$$=f(a)\cos na-f(a)\cos na=0.$$

It follows that the sum (J) composed of pairs of these terms vanishes, whence

$$\lim_{\Delta\to 0} \sum_{0}^{2\pi} f(x) \cos nx\, \Delta x = \int_{0}^{2\pi} f(x) \cos nx\, dx = 0.$$

We now show that

$$\int_{0}^{2\pi} f(x) \sin nx\, dx = 2\int_{0}^{\pi} f(x) \sin nx\, dx.$$

We use the same method to form the sum

$$\sum_{0}^{2\pi} f(x) \sin nx\, \Delta x.$$

Observing that

$$\sin n\,(2\pi - a) = \sin\,(2\pi n - na) = -\sin na,$$

we have

$$f(2\pi - a) \sin n\,(2\pi - a) = [-f(a)]\,(-\sin na) = f(a) \sin na.$$

It follows that, for each term $f(a) \sin na$ that we have for a point of the interval $[0, \pi]$, we have the same term for the corresponding point of the interval $[\pi, 2\pi]$. Hence

$$\sum_{0}^{\pi} f(x) \sin nx\, \Delta x = \sum_{\pi}^{2\pi} f(x) \sin nx\, dx$$

and

$$\sum_{0}^{2\pi} f(x) \sin nx\, \Delta x = 2\sum_{0}^{\pi} f(x) \sin nx\, \Delta x.$$

Thus

$$\lim_{\Delta x\to 0} \sum_{0}^{2\pi} f(x) \sin nx\, \Delta x = \lim_{\Delta x\to 0} \left[2\sum_{0}^{\pi} f(x) \sin nx\, \Delta x\right]$$

$$= 2\int_{0}^{\pi} f(x) \sin nx\, dx.$$

We see from the results obtained that all the coefficients a_n $(n = 0, 1, 2, \ldots)$ of the Fourier series of the function in question

vanish, since

$$a_n = \frac{1}{\pi}\int_0^{2\pi} f(x)\cos nx\,dx = 0. \tag{A}$$

On the other hand, the coefficients b_n of the Fourier series will be given by

$$b_n = \frac{1}{\pi}\int_0^{2\pi} f(x)\sin nx\,dx = \frac{2}{\pi}\int_0^{\pi} f(x)\sin nx\,dx. \tag{B}$$

The Fourier series of a function whose graph is symmetrical about the point $(\pi, 0)$ is therefore of the form

$$b_1 \sin x + b_2 \sin 2x + b_3 \sin 3x + \dots$$

2. We take a function in the segment $[0, 2\pi]$ which is symmetrical with respect to the straight line $x = \pi$ (fig. 121).

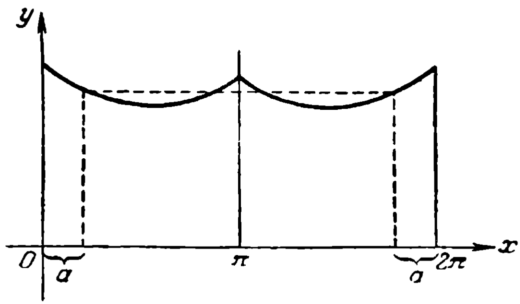


FIG. 121

We have for such a function

$$f(2\pi - a) = f(a).$$

Hence

$$f(2\pi - a)\cos n(2\pi - a) = f(a)\cos na,$$

and

$$f(2\pi - a)\cdot\sin n(2\pi - a) = -f(a)\sin na.$$

We obtain, by the same arguments as in 1.,

$$\int_0^{2\pi} f(x)\cos nx\,dx = 2\int_0^{\pi} f(x)\cos nx\,dx$$

and $$\int_0^{2\pi} f(x)\sin nx\,dx=0.$$

Hence

$$a_n=\frac{1}{\pi}\int_0^{2\pi} f(x)\cos nx\,dx=\frac{2}{\pi}\int_0^{\pi} f(x)\cos nx\,dx \quad (n=0,\ 1,\ 2,\ \dots)$$

and

$$b_n=\frac{1}{\pi}\int_0^{2\pi} f(x)\sin nx\,dx=0 \quad (n=1,\ 2,\ \dots).$$

A function whose graph is symmetrical with respect to the straight line $x=\pi$ thus has a Fourier series of the form

$$\frac{a_0}{2}+a_1\cos x+a_2\cos 2x+a_3\cos 3x+\dots$$

§ 110. Fourier expansions of some functions commonly encountered in electrical work *.

All the present functions are periodic, of period 2π, so that they can be sufficiently defined in the interval $[0, 2\pi]$; they are

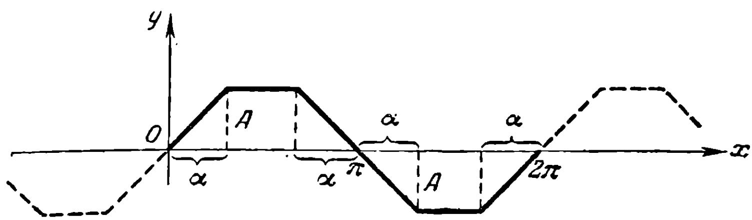


FIG. 122

defined at other points with the aid of the relationship $f(x+2\pi)=f(x)$, i.e. their graphs consist of a repetition of the graph in the interval $[0, 2\pi]$.

* All the functions considered in this article satisfy the conditions of the theorem of 6. § 108,

1. We shall expand the function illustrated in fig. 122 as a Fourier series.

Since the graph is symmetrical about the point $(\pi, 0)$ the expansion will only contain sines, i.e. $a_n=0$ $(n=0, 1, 2, \ldots)$ and

$$b_n=\frac{2}{\pi}\int_0^{\pi} f(x)\sin nx\,dx \quad (n=1, 2, 3, \ldots).$$

It follows from the graph that the function $f(x)$ concerned is defined as follows in the interval $[0, \pi]$:

$$f(x)=\begin{cases} \dfrac{A}{\alpha}x & \text{for } 0 \leqslant x \leqslant \alpha, \\ A & \text{for } \alpha \leqslant x \leqslant \pi-\alpha, \\ -\dfrac{A}{\alpha}(x-\pi) & \text{for } \pi-\alpha \leqslant x \leqslant \pi. \end{cases}$$

Hence

$$b_n=\frac{2}{\pi}\int_0^{\pi} f(x)\sin nx\,dx$$

$$=\frac{2}{\pi}\left\{\frac{A}{\alpha}\int_0^{\alpha} x\sin nx\,dx+A\int_{\alpha}^{\pi-\alpha}\sin nx\,dx\right.$$

$$\left.-\frac{A}{\alpha}\int_{\pi-\alpha}^{\pi}(x-\pi)\sin nx\,dx\right\}$$

or

$$b_n=\frac{2A}{\alpha\pi}\left\{\int_0^{\alpha} x\sin nx\,dx+\alpha\int_{\alpha}^{\pi-\alpha}\sin nx\,dx-\int_{\pi-\alpha}^{\pi}(x-\pi)\sin nx\,dx\right\}.$$

We shall evaluate each integral separately. Integration by parts gives us

$$\int_0^\alpha x \sin nx\,dx = -\frac{1}{n}[x\cos nx]_0^\alpha + \frac{1}{n^2}[\sin nx]_0^\alpha$$

$$= -\frac{1}{n}\cdot\alpha\cdot\cos n\alpha + \frac{1}{n^2}\sin n\alpha.$$

Also,

$$\alpha\int_\alpha^{\pi-\alpha} \sin nx\,dx = -\frac{\alpha}{n}[\cos nx]_\alpha^{\pi-\alpha}$$

$$= -\frac{\alpha}{n}[\cos n(\pi-\alpha) - \cos n\alpha] = -\frac{\alpha}{n}(\cos n\pi\cos n\alpha +$$

$$+\sin n\pi \sin n\alpha - \cos n\alpha) = -\frac{\alpha}{n}(\cos n\alpha\cos n\pi - \cos n\alpha)$$

(since $\sin n\pi = 0$ with $n = 1, 2, \ldots$).

As regards the final integral, we have

$$-\int_{\pi-\alpha}^{\pi}(x-\pi)\sin nx\,dx = \int_{\pi-\alpha}^{\pi}(\pi-x)\sin nx\,dx = I.$$

We use the method of integration by parts, writing

$$\pi - x = u\,, \qquad \sin nx\,dx = dv,$$

in which case

$$du = -dx\,, \qquad v = -\frac{1}{n}\cos nx,$$

so that we obtain

$$I = -\frac{1}{n}[(\pi-x)\cos nx]_{\pi-\alpha}^{\pi} - \frac{1}{n}\int_{\pi-\alpha}^{\pi}\cos nx\,dx$$

$$= -\frac{1}{n}[-\alpha\cos n(\pi-\alpha)] - \frac{1}{n^2}[\sin nx]_{\pi-\alpha}^{\pi}$$

$$=\frac{\alpha}{n}\cos n(\pi-\alpha)-\frac{1}{n^2}[\sin n\pi-\sin n(\pi-\alpha)]$$

or, since $\sin n\pi=0$,

$$I=\frac{1}{n}\Big[\alpha\cos n\pi\cos n\alpha+\alpha\sin n\pi\sin n\alpha+$$

$$+\frac{1}{n}\sin n\pi\cos n\alpha-\frac{1}{n}\cos n\pi\sin n\alpha\Big]$$

$$=\frac{1}{n}\left(\alpha\cos n\pi\cos n\alpha-\frac{1}{n}\cos n\pi\sin n\alpha\right).$$

Hence

$$b_n=\frac{2A}{\alpha\pi}\Big[-\frac{\alpha}{n}\cos n\alpha+\frac{1}{n^2}\sin n\alpha-\frac{\alpha}{n}\cos n\alpha\cos n\pi+$$

$$+\frac{\alpha}{n}\cos n\alpha+\frac{\alpha}{n}\cos n\pi\cos n\alpha-\frac{1}{n^2}\cos n\pi\sin n\alpha\Big]$$

$$=\frac{2A}{\pi\alpha n^2}\sin n\alpha(1-\cos n\pi).$$

We now have

$$b_1=\frac{4A}{\alpha\pi}\sin\alpha\,,\ b_2=0\,,\ b_3=\frac{4A}{9\alpha\pi}\sin 3\alpha\,,\ b_4=0\,,\ b_5=\frac{4A}{25\alpha\pi}\sin 5\alpha\,,$$

$$b_6=0\,,\ b_7=\frac{4A}{49\alpha\pi}\sin 7\alpha\,,\ \ldots$$

so that

$$f(x)=\frac{4A}{\pi\alpha}\Big[\sin\alpha\sin x+\frac{1}{9}\sin 3\alpha\sin 3x+\frac{1}{25}\sin 5\alpha\sin 5x+$$

$$+\frac{1}{49}\sin 7\alpha\sin 7x+\ldots\Big]. \tag{58}$$

Since $f(x)$ is continuous in $[0\,,2\pi]$, the sum $S(x)$ of its Fourier series will coincide with $f(x)$ at all interior points of the interval.

Moreover, since $f(0)=f(2\pi)=0$, whilst the sum $S(x)$ of its Fourier series is equal to the arithmetic mean $\frac{f(0)+f(2\pi)}{2}=0$ at the ends of the interval, we can conclude that the expansion obtained is valid throughout the interval $[0, 2\pi]$.

In particular, with $\alpha=\frac{\pi}{3}$ we have

$$\sin\alpha=\frac{\sqrt{3}}{2}, \quad \sin 3\alpha=0, \quad \sin 5\alpha=-\frac{\sqrt{3}}{3}, \quad \sin 7\alpha=\frac{\sqrt{3}}{2}, \ldots$$

Expansion (58) in this case takes the form

$$f(x)=\frac{6\sqrt{3}A}{\pi^2}\left(\sin x-\frac{1}{25}\sin 5x+\frac{1}{49}\sin 7x-\ldots\right).$$

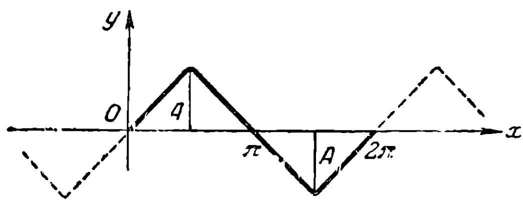


FIG. 123

The trapezium degenerates to a triangle with $\alpha=\frac{1}{2}\pi$ (fig. 123) and the function is defined in the interval $[0, \pi]$ as

$$f(x)=\begin{cases} \frac{2A}{\pi}x & \text{for} \quad 0\leqslant x\leqslant\frac{\pi}{2}, \\ -\frac{2A}{\pi}(x-\pi) & \text{for} \quad \frac{\pi}{2}\leqslant x\leqslant\pi. \end{cases} \tag{59}$$

On using expansion (58) with $\alpha=\frac{1}{2}\pi$ we arrive at the following expansion of function (59) in the interval $[0, 2\pi]$:

$$f(x)=\frac{8A}{\pi^2}\left(\sin x-\frac{1}{9}\sin 3x+\frac{1}{25}\sin 5x-\frac{1}{49}\sin 7x+\ldots\right).$$

2. We find the Fourier expansion in the interval $[0, 2\pi]$ of the function defined as follows:

$$f(x)=\begin{cases} A\sin x & \text{for} \quad 0 \leqq x \leqq \pi, \\ 0 & \text{for} \quad \pi \leqq x \leqq 2\pi \end{cases}$$

(fig. 124). We have

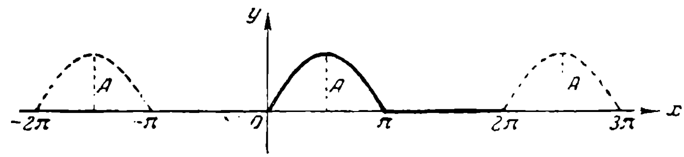


FIG. 124

$$\frac{a_0}{2}=\frac{1}{2\pi}\int_0^{2\pi} f(x)\,dx=\frac{1}{2\pi}\left\{\int_0^{\pi} A\sin x\,dx+\int_0^{2\pi} 0\,.\,dx\right\}$$

$$=\frac{A}{2\pi}\left[-\cos x\right]_0^{\pi}=\frac{A}{\pi}\,.$$

$$a_1=\frac{1}{\pi}\int_0^{2\pi} f(x)\cos x\,dx=\frac{1}{\pi}\left\{\int_0^{\pi} A\sin x\cos x\,dx+\int_{\pi}^{2\pi} 0\,.\,dx\right\}$$

$$=\frac{A}{2\pi}\int_0^{\pi}\sin 2x\,dx=-\frac{1}{4\pi}\left[\cos 2x\right]_0^{\pi}=0.$$

For $n=2, 3, \ldots$ we find

$$a_n=\frac{1}{\pi}\int_0^{2\pi} f(x)\cos nx\,dx=\frac{1}{\pi}\left\{\int_0^{\pi} A\sin x\cos nx\,dx+\int_{\pi}^{2\pi} 0\,.\,dx\right\}$$

$$=\frac{A}{2\pi}\int_0^{\pi}\left[\sin(n+1)x-\sin(n-1)x\right]dx$$

$$=\frac{A}{2\pi}\left\{-\frac{1}{n+1}\left[\cos(n+1)x\right]_0^\pi+\frac{1}{n-1}\left[\cos(n-1)x\right]_0^\pi\right\}$$

$$=\frac{A}{2\pi}\left\{-\frac{1}{n+1}\left[\cos(n+1)\pi-1\right]+\frac{1}{n-1}\left[\cos(n-1)\pi-1\right]\right\}.$$

Hence

$$a_2=\frac{A}{2\pi}\left\{-\frac{1}{3}(\cos 3\pi-1)+(\cos\pi-1)\right\}$$

$$=\frac{A}{2\pi}\left\{\frac{2}{3}-2\right\}=-\frac{2A}{3\pi};$$

$$a_3=\frac{A}{2\pi}\left\{-\frac{1}{4}(\cos 4\pi-1)+\frac{1}{2}(\cos 2\pi-1)\right\}=0;$$

$$a_4=\frac{A}{2\pi}\left\{-\frac{1}{5}(\cos 5\pi-1)+\frac{1}{3}(\cos 3\pi-1)\right\}$$

$$=\frac{A}{2\pi}\left(\frac{2}{5}-\frac{2}{3}\right)=-\frac{2A}{15\pi};$$

$$a_5=0;$$

$$a_6=\frac{A}{2\pi}\left\{-\frac{1}{7}(\cos 7\pi-1)+\frac{1}{5}(\cos 5\pi-1)\right\}=-\frac{2A}{35\pi};$$

Again

$$b_1=\frac{1}{\pi}\int_0^{2\pi} f(x)\sin x\,dx=\frac{1}{\pi}\left\{\int_0^{\pi} A\sin^2 x\,dx+\int_\pi^{2\pi} 0\,dx\right\}$$

$$=\frac{A}{2\pi}\int_0^{\pi}(1-\cos 2x)\,dx=\frac{A}{2\pi}\left\{[x]_0^\pi-\frac{1}{2}[\sin 2x]_0^\pi\right\}=\frac{A}{2}.$$

For $n=2, 3, \ldots$ we find

$$b_n=\frac{1}{\pi}\int_0^{\pi} A\sin x\sin nx\,dx=\frac{A}{2\pi}\int_0^{\pi}[\cos(n-1)x-\cos(n+1)x]\,dx$$

$$=\frac{A}{2\pi}\left\{\frac{1}{n-1}[\sin(n-1)x]_0^{\pi}-\frac{1}{n+1}[\sin(n+1)x]_0^{\pi}\right\}=0.$$

The required expansion is therefore

$$f(x)=\frac{A}{\pi}+\frac{A}{2}\sin x-\frac{2A}{3\pi}\cos 2x-\frac{2A}{15\pi}\cos 4x-\dots$$

$$=\frac{2A}{\pi}\left(\frac{1}{2}+\frac{\pi}{4}\sin x-\frac{\cos 2x}{3}-\frac{\cos 4x}{15}-\frac{\cos 6x}{35}-\right)\dots$$

3. We shall expand the function

$$f(x)=A\,|\sin x|$$

as a Fourier series (fig. 125). Since the function is symmetrical with respect to the straight line $x=\pi$, its expansion will contain cosines only, while

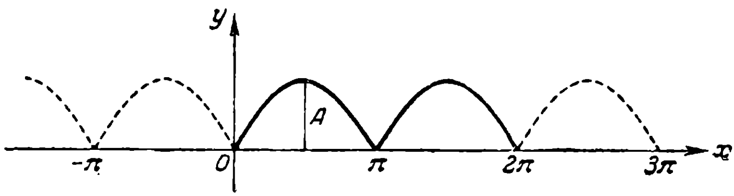


FIG. 125

$$a_n=\frac{2}{\pi}\int_0^{\pi} f(x)\cos nx\,dx=\frac{2A}{\pi}\int_0^{\pi}\sin x\cos nx\,dx.$$

The coefficients $a_0, a_1, a_2, \dots$ are thus evaluated as in the previous example and differ from these only by a factor of 2, while $b_1=0$. We thus obtain

$$A\,|\sin x|=\frac{2A}{\pi}\left(1-\frac{2}{3}\cos 2x-\frac{2}{15}\cos 4x-\frac{2}{35}\cos 6x-\dots\right).$$

4. We find the Fourier expansion of the function defined as follows in the interval $[0, 2\pi]$:

$$f(x)=\begin{cases} A & \text{for} \quad 0\leqslant x\leqslant\pi, \\ -A & \text{for} \quad \pi<x\leqslant 2\pi \end{cases} \qquad (A>0)$$

(fig. 126). This function is symmetrical about the point $(\pi, 0)$ (in accordance with what was said in 4. § 108, we associate with

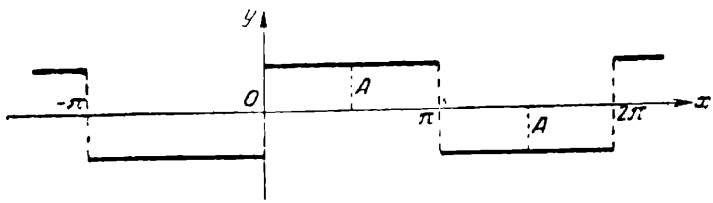


FIG. 126

the point (π, A) the symmetrical point $(\pi, -A)$). The Fourier expansion will thus contain sines only. We evaluate the Fourier coefficients b_n in accordance with formula (B) of § 109, whence

$$b_n=\frac{2}{\pi}\int_0^{\pi} A\sin nx\,dx=\frac{2A}{\pi}\int_0^{\pi}\sin nx\,dx$$

$$=-\frac{2A}{\pi n}[\cos nx]_0^{\pi}=-\frac{2A}{\pi n}(\cos n\pi-1).$$

We find

$$b_1=\frac{4A}{\pi}, \quad b_2=0, \quad b_3=\frac{4A}{3\pi},$$

$$b_4=0, \quad b_5=\frac{4A}{5\pi}, \quad b_6=0, \quad b_7=\frac{4A}{7\pi}, \ldots$$

Consequently,

$$f(x)=\frac{4A}{\pi}\left(\sin x+\frac{1}{3}\sin 3x+\frac{1}{5}\sin 5x+\frac{1}{7}\sin 7x+\ldots\right).$$

Fig. 127 illustrates the graph of the function for the case $A = \frac{\pi}{4}$ and the graphs of consecutive partial sums of its Fourier series, i.e. of the series

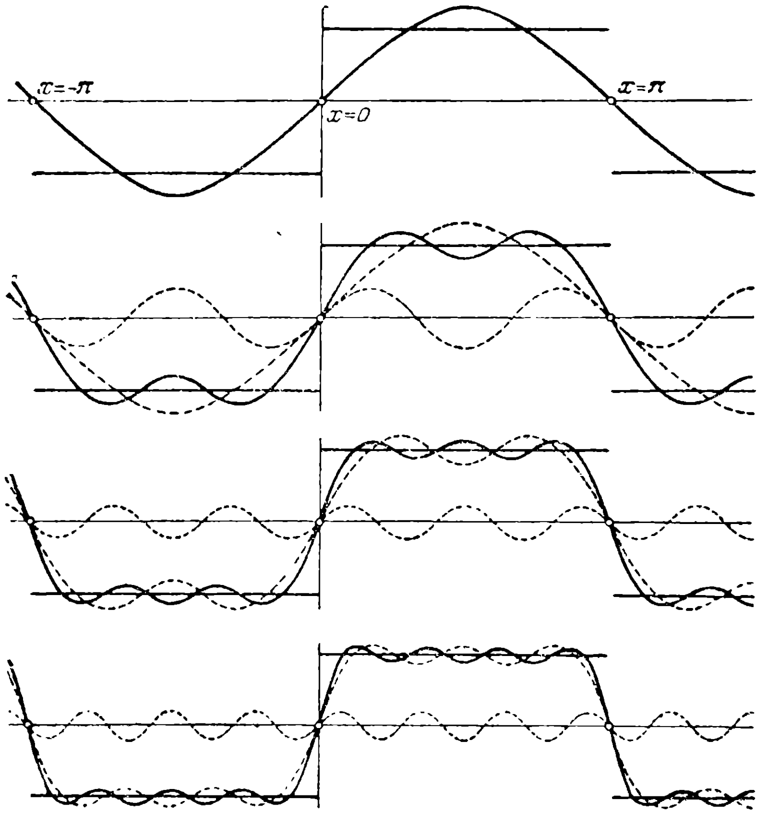


FIG. 127

$$\sin x + \frac{1}{3}\sin 3x + \frac{1}{5}\sin 5x + \frac{1}{7}\sin 7x, \ldots$$

The required function is drawn in each sketch. The first sketch contains, in addition, the partial sum $S_1(x) = \sin x$. The second shows $y = \sin x$ and $y = \frac{1}{3}\sin 3x$ as dotted curves and the partial

sum $S_2(x) = \sin x + \frac{1}{3}\sin 3x$ as a full curve, obtained by adding the simple periodic curves of $\sin x$ and $\frac{1}{3}\sin 3x$. The third sketch gives $S_2(x)$, now shown as a dotted curve, the curve $y = \frac{1}{5}\sin 5x$ (dotted) and the partial sum $S_3(x) = S_2(x) + \frac{1}{5}\sin 5x = \sin x + \frac{1}{3}\sin 3x + \frac{1}{5}\sin 5x$, and so on.

It is clear from the figure how the sums of elementary periodic functions approximate more and more closely, as indicated in § 107, to a periodic function which is not itself elementary.

EXERCISES

On § 96

Establish the convergence or divergence of the following series with the aid of the comparison test:

1. $1 + \frac{2}{3} + \frac{2^2}{3.5} + \frac{2^3}{3.5.7} + \frac{2^4}{3.5.7.9} + \ldots$ Ans. Convergent.

2. $\frac{3}{2} + \frac{4}{3.2} + \frac{5}{4.3} + \frac{6}{5.6} + \ldots + \frac{n+2}{(n+1)\,n}$ Ans. Divergent

3. $1 + \frac{1}{3^2} + \frac{1}{5^2} + \ldots + \frac{1}{(2n-1)^2} + \ldots$ Ans. Convergent.

4. $1 + \frac{1}{5} + \frac{1}{10} + \frac{1}{17} + \ldots + \frac{1}{n^2+1} + \ldots$ Ans. Convergent.

5. $\frac{1}{\ln 2} + \frac{1}{\ln 3} + \frac{1}{\ln 4} + \ldots + \frac{1}{\ln (n+1)} + \ldots$ Ans. Divergent.

Investigate with d'Alembert's test the convergence of the series:

6. $\frac{1}{2.1} + \frac{1}{2^3.3} + \frac{1}{2^5.5} + \ldots + \frac{1}{2^{2n-1}(2n-1)} + \ldots$ Ans. Convergent.

7. $1+\frac{5}{2!}+\frac{5^2}{3!}+\ldots+\frac{5^{n-1}}{n!}+\ldots$ Ans. Convergent.

8. $\frac{2}{1.2}+\frac{2^2}{2.3}+\frac{2^3}{3.4}+\ldots+\frac{2_n}{n(n+1)}+\ldots$ Ans. Divergent.

9. $\frac{1}{3}+\frac{2}{3^2}+\frac{3}{3^3}+\ldots+\frac{n}{3^n}+\ldots$ Ans. Convergent.

10. $1+\frac{2^2}{3}+\frac{3^2}{3^2}+\frac{4^2}{3^3}+\ldots+\frac{n^2}{3^{n-1}}+\ldots$ Ans. Convergent.

11. $1+\frac{3}{2.3}+\frac{3^2}{2^2.5}+\frac{3^3}{2^3.7}+\ldots+\frac{3^{n-1}}{2^{n-1}(2n-1)}+\ldots$ Ans. Divergent

On §§ 97 *and* 98

Establish whether the following series are absolutely or non-absolutely convergent.

12. $1-\frac{1}{3^2}+\frac{1}{5^2}-\frac{1}{7^2}+\frac{1}{9^2}-\ldots$ Ans. Absolutely convergent.

13. $\frac{1}{2}-\frac{1}{2}\cdot\frac{1}{2^2}+\frac{1}{3}\cdot\frac{1}{2^3}-\frac{1}{4}\cdot\frac{1}{2^4}+\ldots$ Ans. Absolutely convergent.

14. $\frac{1}{\ln 2}-\frac{1}{\ln 3}+\frac{1}{\ln 4}-\ldots$ Ans. Non-absolutely convergent.

15. $-1+\frac{1}{\sqrt{2}}-\frac{1}{\sqrt{3}}+\frac{1}{\sqrt{4}}-\ldots$ Ans. Non-absolutely convergent.

16. $\frac{\sin\alpha}{1}+\frac{\sin 2\alpha}{4}+\frac{\sin 3\alpha}{9}+\ldots$ Ans. Absolutely convergent.

17. $1-\frac{1}{3}+\frac{1}{5}-\frac{1}{7}+\frac{1}{9}-\ldots$ Ans. Non-absolutely convergent.

18. $\frac{1}{3}-\left(\frac{2}{5}\right)^2+\left(\frac{3}{7}\right)^3-\left(\frac{4}{9}\right)^4+\ldots$ Ans. Absolutely convergent.

On § 100

Find the intervals of convergence of the following series and decide whether or not they are convergent at the ends of this interval.

19. $x+\frac{x^3}{3}+\frac{x^5}{5}+\ldots+\frac{x^{2n-1}}{2n-1}+\ldots$ Ans. $-1<x<1$.

20. $\frac{x}{2^2}+\frac{x^2}{4^2}+\frac{x^3}{6^2}+\ldots$ Ans. $-1\leqslant x\leqslant 1$.

21. $\frac{x}{1\,.\,2}+\frac{x^2}{3\,.\,4}+\frac{x^3}{5\,.\,6}+\ldots$ Ans. $-1\leqslant x\leqslant 1$.

22. $\frac{1}{2}+\frac{x}{2^2}+\frac{x^2}{2^3}+\ldots$ Ans. $-2<x<2$.

23. $1+\frac{x}{3\,.\,2}+\frac{x^2}{3^2\,.\,3}+\frac{x^3}{3^3\,.\,4}+\ldots$ Ans. $-3\leqslant x<3$.

24. $1-\frac{x}{5\sqrt{2}}+\frac{x^2}{5^2\sqrt{3}}-\frac{x^3}{5^3\sqrt{4}}+\ldots$ Ans. $-5<x\leqslant 5$.

On §§ 102, 103 *and* 104

Use the Maclaurin expansions of the functions e^z, $\sin z$, $\cos z$, $\ln(1+z)$ and $(1+z)^\alpha$ to expand the following functions as power series:

25. e^{2x}. 26. e^{-x^2}. 27. $y=\sin\frac{x}{2}$.

28. $y=\cos^2 x$.

Ans. $1-\left[x^2-\frac{2^3x^4}{4!}+\ldots+(-1)^{n-1}\frac{2^{2n-1}x^{2n}}{(2n)!}+\ldots\right]$.

29. $\frac{1}{\sqrt[3]{1+x^3}}$.

Ans. $1-\left[\frac{1}{3}x^3-\frac{1\,.\,4}{3\,.\,6}x^6+\ldots+(-1)^{n+1}\frac{1\,.\,4\ldots(3n-2)}{3^n\,.\,n!}x^{3n}+\ldots\right]$.

30. $\ln \frac{1+x}{1-x}$. Ans. $2\left(x+\frac{x^3}{3}+\frac{x^5}{5}+\ldots+\frac{x^{2n-1}}{2n-1}+\ldots\right)$.

Hint. Use the formula $\ln \frac{1+x}{1-x} = \ln(1+x) - \ln(1-x)$ and obtain the required series by term by term subtraction of the series for $\ln(1+x)$ and $\ln(1-x)$.

31. $\frac{e^x+e^{-x}}{2}$. Ans. $1+\frac{x^2}{2!}+\ldots+\frac{x^{2n-2}}{(2n-2)!}+\ldots$

On §§ 108 *and* 109

Expand the following functions as Fourier series (in the interval $[0, 2\pi]$):

32. $f(x)=\begin{cases} -x & \text{for } 0 \leqslant x \leqslant \pi, \\ 0 & \text{for } \pi < x \leqslant 2\pi. \end{cases}$

33. $f(x)=x^2$.

34. $f(x)=\begin{cases} 0 & \text{for } 0 \leqslant x \leqslant \pi, \\ x^2 & \text{for } \pi < x \leqslant 2\pi \end{cases}$